GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 02:35:54 ; Search time 3265 Seconds
(without alignments)
12136.941 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtggtcacaggtag.....aaaaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	751	56.6	835	14	CA310925	CA310925 UI-CF-FN0
c	2	721	54.3	769	14	CA444761	CA444761 UI-H-DH1-
	3	688	51.8	1049	12	BM457033	BM457033 AGENCOURT
	4	656	49.4	781	14	CB956370	CB956370 AGENCOURT
c	5	634	47.8	634	12	BQ018441	BQ018441 UI-H-DH1-
c	6	632	47.6	632	12	BM994555	BM994555 UI-H-DH0-
c	7	587	44.2	682	14	CA309985	CA309985 UI-H-FT1-
c	8	582	43.9	633	12	BM994326	BM994326 UI-H-DH0-
	9	519	39.1	519	9	AL138362	AL138362 DKFZp762F
	10	496	37.4	510	13	BX283437	BX283437 BX283437
c	11	489	36.9	501	9	AL138363	AL138363 DKFZp762F
c	12	455	34.3	470	10	AW515028	AW515028 xu90e05.x
c	13	440	33.2	486	9	AI457621	AI457621 tj63d11.x
c	14	434	32.7	451	14	CA434008	CA434008 UI-H-DF0-
c	15	431	32.5	477	10	BE671191	BE671191 7e55b02.x
c	16	423	31.9	557	9	AU154016	AU154016 AU154016
c	17	421	31.7	423	9	AI066431	AI066431 oq98e10.x
	18	421	31.7	501	13	BQ187235	BQ187235 UI-E-EJ1-
	19	415	31.3	983	10	BE795637	BE795637 601590620
c	20	405	30.5	456	10	BE856182	BE856182 7f89g01.x
c	21	403	30.4	456	9	AA700789	AA700789 zi66c04.s
c	22	399	30.1	502	14	CK300567	CK300567 UI-E-EJ1-
c	23	391	29.5	512	9	AI681374	AI681374 tx46a12.x
c	24	378	28.5	431	14	R71654	R71654 yi53g05.s1
	25	370	27.9	497	10	BE244684	BE244684 TCBAP2E05
c	26	319	24.0	424	14	CA434086	CA434086 UI-H-DF0-
	27	305	23.0	543	14	R71157	R71157 yi53g05.r1
c	28	304	22.9	304	9	AI383065	AI383065 tc20g06.x
c	29	299	22.5	376	10	BE245533	BE245533 TCBAP1D32
	30	286	21.6	385	9	AA644653	AA644653 af74f11.r
c	31	285	21.5	434	10	AW139822	AW139822 UI-H-BI1-
	32	283	21.3	431	14	H47130	H47130 yp73a07.r1
	33	283	21.3	440	14	H78426	H78426 yu22c04.r1
	34	280	21.1	513	9	AA037876	AA037876 zf04f10.r
c	35	277	20.9	543	9	AW070327	AW070327 xa10c02.x
	36	269	20.3	389	14	T85444	T85444 yd82c03.r1
c	37	269	20.3	470	10	BE219390	BE219390 hv57c03.x
c	38	260	19.6	337	10	BE245721	BE245721 TCBAP1D19
	39	255	19.2	749	10	BF696473	BF696473 602125145
c	40	247	18.6	270	13	BU754096	BU754096 UI-1-BB1p
	41	241	18.2	378	10	AW964518	AW964518 EST376591
c	42	226	17.0	350	9	AA191674	AA191674 zq43g04.s
c	43	225	17.0	334	10	BE817016	BE817016 RC0-BN024
	44	221	16.7	405	12	BG912772	BG912772 602807993
	45	221	16.7	768	9	AL554421	AL554421 AL554421
	46	212	16.0	931	13	BU527114	BU527114 AGENCOURT
c	47	211	15.9	310	9	AI864353	AI864353 wg78f06.x
c	48	196	14.8	270	12	BQ025918	BQ025918 UI-1-BB1p
c	49	194	14.6	488	9	AA037877	AA037877 zf04f10.s
c	50	182	13.7	401	12	BM678923	BM678923 UI-E-EO0-
	51	181	13.6	295	13	BQ367558	BQ367558 PM4-GN030
c	52	180	13.6	346	9	AA909074	AA909074 om32a02.s
c	53	164	12.4	342	9	AI242830	AI242830 qi28c12.x
c	54	163	12.3	295	14	H78226	H78226 yu22c04.s1
	55	153	11.5	363	9	AA320033	AA320033 EST22378
	56	150	11.3	1129	12	BM547517	BM547517 AGENCOURT
c	57	138	10.4	153	12	BQ022505	BQ022505 UI-1-BB1p
	58	124	9.3	269	10	BF841308	BF841308 RC2-HT107
	59	113	8.5	304	10	BE246494	BE246494 TCBAP1E47
	60	110	8.3	593	9	AV705109	AV705109 AV705109
c	61	105	7.9	135	12	BQ022575	BQ022575 UI-1-BB1p
	62	100	7.5	901	13	BX348422	BX348422 BX348422
c	63	84	6.3	334	12	BQ025276	BQ025276 UI-1-BB1p
c	64	74	5.6	186	14	Z41723	Z41723 HSC2BA022 n
	65	70	5.3	906	13	BU171407	BU171407 AGENCOURT
	66	67	5.0	341	10	BF987119	BF987119 MR2-GN012
c	67	62	4.7	480	14	T85347	T85347 yd82c03.s1
c	68	43	3.2	732	9	AL554364	AL554364 AL554364
	69	41	3.1	432	14	CF117086	CF117086 fp220.z1
	70	41	3.1	551	10	BE758068	BE758068 212688 MA
	71	41	3.1	552	10	BE753612	BE753612 206620 MA
	72	35	2.6	300	14	Z44845	Z44845 HSC2BA021 n
	73	34	2.6	394	12	BI899337	BI899337 481396 MA
	74	31	2.3	550	10	BE753490	BE753490 206539 MA
	75	30	2.3	312	14	CF116526	CF116526 ad344.z1
	76	30	2.3	524	14	CF116527	CF116527 ad344.z2
	77	29	2.2	167	14	CF313132	CF313132 HD--01-C1

c	78	28	2.1	154	13	BQ792699
	79	28	2.1	171	12	BG688740
c	80	28	2.1	199	10	BF463886
	81	28	2.1	234	13	BU954379
c	82	28	2.1	342	12	BI300710
	83	28	2.1	351	9	AA184228
c	84	28	2.1	358	12	BI300041
c	85	28	2.1	406	12	BI297552
c	86	28	2.1	464	12	BI298701
	87	28	2.1	470	10	BF776486
c	88	28	2.1	478	10	BF457922
c	89	28	2.1	488	12	BG373733
c	90	28	2.1	518	10	BE996205
c	91	28	2.1	519	9	AI227813
c	92	28	2.1	560	12	BM900146
c	93	28	2.1	623	12	BI300357
	94	28	2.1	652	28	BH684258
	95	28	2.1	774	29	CC536668
	96	28	2.1	1086	14	CK231937
c	97	27	2.0	72	10	AW169355
c	98	27	2.0	117	14	CF201671
	99	27	2.0	154	12	BI003499
c	100	27	2.0	209	14	CB821576
c	101	27	2.0	210	12	BI845801
	102	27	2.0	229	12	BI845510
	103	27	2.0	346	10	BF191229
c	104	27	2.0	349	10	BE674222
c	105	27	2.0	353	13	BQ172169
	106	27	2.0	358	10	BF191683
c	107	27	2.0	369	9	AA280737
c	108	27	2.0	372	10	AW138882
	109	27	2.0	396	12	BM027121
	110	27	2.0	405	14	CA336437
c	111	27	2.0	440	29	CE533755
c	112	27	2.0	448	12	BM874938
	113	27	2.0	477	10	BF729120
c	114	27	2.0	487	14	CA663515
c	115	27	2.0	493	10	AW129612
	116	27	2.0	501	14	CB129334
c	117	27	2.0	566	10	AW118907
c	118	27	2.0	587	9	AI676005
c	119	27	2.0	587	14	CA814536
c	120	27	2.0	607	14	CB343776
	121	27	2.0	636	29	CE637121
	122	27	2.0	710	10	BF540944
c	123	27	2.0	711	14	CA759041
c	124	27	2.0	713	14	CD373920
	125	27	2.0	714	9	AU068118
c	126	27	2.0	768	14	CB343742
	127	27	2.0	1026	13	BX347228
c	128	27	2.0	1044	14	CD502799
c	129	26	2.0	54	14	CF327511
	130	26	2.0	96	14	CF327512
c	131	26	2.0	101	13	BQ794536
	132	26	2.0	115	14	CF314604
c	133	26	2.0	118	14	CF204241
c	134	26	2.0	164	10	BE946330
	135	26	2.0	189	12	BI941630
c	136	26	2.0	191	14	CF328196
	137	26	2.0	217	12	BI798001
	138	26	2.0	224	14	CB316674
	139	26	2.0	233	12	BI503458
	140	26	2.0	245	14	CF575205
c	141	26	2.0	275	12	BI134712
c	142	26	2.0	279	9	AI611640
	143	26	2.0	297	9	AW021059
c	144	26	2.0	310	9	AA598544
c	145	26	2.0	321	10	AW232024
	146	26	2.0	321	14	H57960
	147	26	2.0	329	14	N75931
c	148	26	2.0	348	12	BG377618
	149	26	2.0	351	10	BE123282
c	150	26	2.0	364	10	BF729569

BQ792699 EST 8419
BG688740 336481 BA
BF463886 UI-M-CG0p
BU954379 AGENCOURT
BI300710 UI-R-CV2-
AA184228 mt32g04.r
BI300041 UI-R-CV2-
BI297552 UI-R-CV2-
BI298701 UI-R-CV2-
BF776486 287298 MA
BF457922 UI-M-BZ1-
BG373733 UI-R-CV1-
BE996205 UI-M-BZ1-
AI227813 EST224508
BM900146 UI-M-DJ1-
BI300357 UI-R-CV2-
BH684258 BOMMN30TF
CC536668 CH240_415
CK231937 ILLUMIGEN
AW169355 xj22g08.x
CF201671 RR890915N
BI003499 MR3-HN012
CB821576 EST 2427
BI845801 fs98b02.x
BI845510 fs98b02.y
BF191229 237997 MA
BE674222 7d77a06.x
BQ172169 WHE2003_A
BF191683 239337 MA
AA280737 zs96h10.s
AW138882 UI-H-BI1-
BM027121 GIT000041
CA336437 NISC_lv01
CE533755 tigr-gss-
BM874938 if39d08.x
BF729120 1000072F0
CA663515 wlmk1.pk0
AW129612 xe19g11.x
CB129334 K-EST0178
AW118907 xd98b11.x
AI676005 wc04b01.x
CA814536 CA48LN10I
CB343776 CA32EN000
CE637121 tigr-gss-
BF540944 602067682
CA759041 BR060014B
CD373920 WHE2613_E
AU068118 AU068118
CB343742 CA32EN000
BX347228 BX347228
CD502799 CDA58-F12
CF327511 NACL--02-
CF327512 NACL--02-
BQ794536 EST 3474
CF314604 HD--03-D0
CF204241 RR890915I
BE946330 UI-M-BZ0-
BI941630 sc80f06.y
CF328196 NACL--02-
BI798001 H094H04 E
CB316674 AGENCOURT
BI503458 BB170031B
CF575205 MCSA144D0
BI134712 UI-M-BH3-
AI611640 tt36h10.x
AW021059 df18c08.y
AA598544 ae39b01.s
AW232024 fj12c10.x
H57960 yr05a02.s1
N75931 za42a01.s1
BG377618 UI-R-CU0-
BE123282 945040D09
BF729569 1000079F0

ALIGNMENTS

RESULT 1
CA310925/c
LOCUS CA310925 835 bp mRNA linear EST 04-NOV-2002
DEFINITION UI-CF-FN0-afb-j-06-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone UI-CF-FN0-afb-j-06-0-UI 3', mRNA sequence.
ACCESSION CA310925
VERSION CA310925.1 GI:24529023
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES Location/Qualifiers
source 1. .835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FN0-afb-j-06-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FN0"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-FN0 is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h
TAG_LIB=UI-CF-FN0
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 56.6%; Score 751; DB 14; Length 835;
Best Local Similarity 100.0%; Pred. No. 2.8e-123;
Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 575 ATTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCC 634
Db 751 ATTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCC 692

Qy 635 CATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCC 694

```
Db        691 CATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCC 632

Qy        695 ACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTC 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 ACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTC 572

Qy        755 CCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCT 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCT 512

Qy        815 TGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGC 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 TGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGC 452

Qy        875 AAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 AAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGG 392

Qy        935 AGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGG 332

Qy        995 AGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTC 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 AGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTC 272

Qy       1055 CTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGG 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGG 212

Qy       1115 AAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGG 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 AAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGG 152

Qy       1175 CAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGC 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         151 CAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGC 92

Qy       1235 AGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAA 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          91 AGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAA 32

Qy       1295 AGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1325
              ||||||||||||||||||||||||||||||
Db          31 AGTTACAGATGTCAAAAAAAAAAAAAAAAA 1


RESULT 2
CA444761/c
LOCUS       CA444761                 769 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-DH1-awv-f-12-0-UI.s1 NCI_CGAP_DH1 Homo sapiens cDNA clone
            UI-H-DH1-awv-f-12-0-UI 3', mRNA sequence.
ACCESSION   CA444761
VERSION     CA444761.1  GI:24809181
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-DH1-awv-f-12-0-UI"
                     /tissue_type="Metastatic Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DH1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DH1 is a normalized cDNA library containing the
                     following tissue(s): VS-8 Cell line from Metastatic
                     Chondrosarcoma in Lung. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is AGATCATTGC.
                     TAG_TISSUE=lung
                     TAG_LIB=UI-H-DH1
                     TAG_SEQ=AGATCATTGC"
ORIGIN

  Query Match           54.3%;  Score 721;  DB 14;  Length 769;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-118;
  Matches  721;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        583 GAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCA 677

Qy        643 GGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCG 617

Qy        703 TGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCT 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCT 557

Qy        763 AAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCA 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 AAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCA 497

Qy        823 AGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCA 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCA 437

Qy        883 CCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACA 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 CCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACA 377

Qy        943 TCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGG 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 TCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGG 317

Qy       1003 GCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCC 257

Qy       1063 TAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTA 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTA 197

Qy       1123 AACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGA 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGA 137
```

```
Qy       1183 ATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTT 1242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 ATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTT 77

Qy       1243 TATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAG 1302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 TATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAG 17

Qy       1303 A 1303
              |
Db          16 A 16
```

RESULT 3
BM457033
LOCUS BM457033 1049 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6406661 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583269 5', mRNA sequence.
ACCESSION BM457033
VERSION BM457033.1 GI:18506073
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1049)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12346 row: b column: 06
High quality sequence stop: 671.

FEATURES Location/Qualifiers
source 1. .1049
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5583269"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 51.8%; Score 688; DB 12; Length 1049;
Best Local Similarity 100.0%; Pred. No. 2.4e-112;
Matches 688; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 174

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 234

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 294

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 354

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 414

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 474

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 534

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 594

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 654

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 714

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 774

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCC 688
              ||||||||||||||||||||||||||||
Db        775 AAGCTGCTGACTGCAGAACTAGGCTCCC 802
```

RESULT 4
CB956370
LOCUS CB956370 781 bp mRNA linear EST 29-APR-2003
DEFINITION AGENCOURT_13666719 NIH_MGC_184 Homo sapiens cDNA clone IMAGE:30353395 5', mRNA sequence.
ACCESSION CB956370
VERSION CB956370.1 GI:30212487
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM154 row: a column: 20
High quality sequence stop: 582.

FEATURES Location/Qualifiers
source 1. .781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30353395"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_184"
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site_1:

SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyoid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.4%; Score 656; DB 14; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.3e-106;
Matches 656; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         84 GGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGTGGCCTGTGGAAGGCA 143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGTGGCCTGTGGAAGGCA 62

Qy        144 CCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGA 203
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGA 122

Qy        204 GGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTGC 263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTGC 182

Qy        264 AGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAAC 323
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAAC 242

Qy        324 CCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTT 383
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTT 302

Qy        384 GTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCTAGACCCCTTGTCAGACTTGGCAT 443
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCTAGACCCCTTGTCAGACTTGGCAT 362

Qy        444 CTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGA 503
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGA 422

Qy        504 CAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCA 563
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCA 482

Qy        564 GACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGG 623
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGG 542

Qy        624 AAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGG 683
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 AAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGG 602

Qy        684 CTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCC 739
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCC 658
```

RESULT 5
BQ018441/c
LOCUS BQ018441 634 bp mRNA linear EST 17-JUN-2002
DEFINITION UI-H-DH1-awq-p-19-0-UI.s1 NCI_CGAP_DH1 Homo sapiens cDNA clone IMAGE:5893122 3', mRNA sequence.
ACCESSION BQ018441
VERSION BQ018441.1 GI:19753718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5893122"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_DH1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_DH1 is a normalized cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATTGC.
TAG_TISSUE=lung
TAG_LIB=UI-H-DH1
TAG_SEQ=AGATCATTGC"

ORIGIN

Query Match 47.8%; Score 634; DB 12; Length 634;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        692 GCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAG 751
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 GCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAG 575

Qy        752 TTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAA 811
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 TTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAA 515

Qy        812 CCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGG 871
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 CCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGG 455

Qy        872 AGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGG 931
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 AGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGG 395

Qy        932 TGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAAC 991
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 TGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAAC 335

Qy        992 TGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTA 1051
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 TGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTA 275
```

```
Qy       1052 GTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGC 1111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGC 215

Qy       1112 CGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTT 1171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         214 CGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTT 155

Qy       1172 GGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCC 1231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         154 GGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCC 95

Qy       1232 AGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAAT 1291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          94 AGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAAT 35

Qy       1292 TAAAGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1325
             ||||||||||||||||||||||||||||||||||
Db          34 TAAAGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 6
BM994555/c
LOCUS       BM994555                 632 bp    mRNA    linear   EST 17-JUN-2002
DEFINITION  UI-H-DH0-aum-m-09-0-UI.s1 NCI_CGAP_DH0 Homo sapiens cDNA clone
            IMAGE:5871536 3', mRNA sequence.
ACCESSION   BM994555
VERSION     BM994555.1  GI:19719456
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 632)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .632
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5871536"
                     /tissue_type="Metastatic Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DH0"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DH0 is a cDNA library containing the following
                     tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma
                     in Lung. The library was constructed according to Bonaldo,
                     Lennon and Soares, Genome Research, 6:791-806, 1996. First
                     strand cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site. Double stranded cDNA was ligated
                     to an EcoR I adaptor, digested with Not I, and cloned
                     directionally into pT7T3-Pac vector. The oligonucleotide
                     used to prime the synthesis of first-strand cDNA contains
                     a library tag sequence that is located between the Not I
                     site and the (dT)18 tail. The sequence tag for this
                     library is AGATCATTGC.
                     TAG_TISSUE=lung
                     TAG_LIB=UI-H-DH0
                     TAG_SEQ=AGATCATTGC"
ORIGIN
```

```
  Query Match             47.6%;  Score 632;  DB 12;  Length 632;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-102;
  Matches  632;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        694 CACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTT 753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 CACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTT 573

Qy        754 CCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACC 813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACC 513

Qy        814 TTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAG 873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAG 453

Qy        874 CAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTG 933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 CAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTG 393

Qy        934 GAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTG 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTG 333

Qy        994 GAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGT 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGT 273

Qy       1054 CCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCG 1113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 CCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCG 213

Qy       1114 GAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGG 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGG 153

Qy       1174 GCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAG 93

Qy       1234 CAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTA 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          92 CAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTA 33

Qy       1294 AAGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1325
              ||||||||||||||||||||||||||||||||
Db          32 AAGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 7
CA309985/c
LOCUS       CA309985                 682 bp    mRNA    linear   EST 01-NOV-2002
DEFINITION  UI-H-FT1-bid-j-07-0-UI.s1 NCI_CGAP_FT1 Homo sapiens cDNA clone
            UI-H-FT1-bid-j-07-0-UI 3', mRNA sequence.
ACCESSION   CA309985
VERSION     CA309985.1  GI:24473039
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES Location/Qualifiers
source 1. .682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT1-bid-j-07-0-UI"
/tissue_type="Aveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FT1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_FT1 is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCCATGCCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
TAG_TISSUE=Human Lung Aveolar Macrophage
TAG_LIB=UI-H-FT1
TAG_SEQ=GGCCATGCCG"

ORIGIN

Query Match 44.2%; Score 587; DB 14; Length 682;
Best Local Similarity 100.0%; Pred. No. 1.8e-94;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       722 TGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTC 781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       607 TGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTC 548

Qy       782 TGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTC 841
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       547 TGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTC 488

Qy       842 ATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACA 901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       487 ATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACA 428

Qy       902 GCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCC 961
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       427 GCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCC 368

Qy       962 TTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAA 1021
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       367 TTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAA 308

Qy      1022 ATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGT 1081
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       307 ATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGT 248

Qy      1082 CTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAA 1141
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       247 CTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAA 188

Qy      1142 ATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAA 1201
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       187 ATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAA 128

Qy      1202 CCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCAT 1261
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       127 CCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCAT 68

Qy      1262 TTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db        67 TTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 21
```

RESULT 8
BM994326/c
LOCUS BM994326 633 bp mRNA linear EST 17-JUN-2002
DEFINITION UI-H-DH0-aul-k-14-0-UI.s1 NCI_CGAP_DH0 Homo sapiens cDNA clone IMAGE:5871109 3', mRNA sequence.
ACCESSION BM994326
VERSION BM994326.1 GI:19719227
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES Location/Qualifiers
source 1. .633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5871109"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_DH0"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_DH0 is a cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATTGC.
TAG_TISSUE=lung
TAG_LIB=UI-H-DH0
TAG_SEQ=AGATCATTGC"

ORIGIN

Query Match 43.9%; Score 582; DB 12; Length 633;
Best Local Similarity 99.8%; Pred. No. 1.5e-93;
Matches 632; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       693 CCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGT 752
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        633 CCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGT 574

Qy        753 TCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAAC 812
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        573 TCCCAAATCTAAATTTTTACAAGAGATTCTGTNTGGGGGAACTTAAGTCAGATCCAGAAC 514

Qy        813 CTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGA 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGA 454

Qy        873 GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGT 932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGT 394

Qy        933 GGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACT 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 GGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACT 334

Qy        993 GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAG 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAG 274

Qy       1053 TCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCC 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 TCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCC 214

Qy       1113 GGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTG 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTG 154

Qy       1173 GGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCA 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCA 94

Qy       1233 GCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATT 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATT 34

Qy       1293 AAAGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1325
              |||||||||||||||||||||||||||||||||
Db         33 AAAGTTACAGATGTCAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 9
AL138362
LOCUS       AL138362                519 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp762F1316_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762F1316_5', mRNA sequence.
ACCESSION   AL138362
VERSION     AL138362.1  GI:6855043
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp762F1316) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES              Location/Qualifiers
     source           1. .519
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="DKFZp762F1316"
                      /tissue_type="melanoma (MeWo cell line)"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="762 (synonym: hmel2)"
                      /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

  Query Match             39.1%;  Score 519;  DB 9;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 2e-82;
  Matches  519;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        364 GGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCTAGA 423
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCTAGA 60

Qy        424 CCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGCCAC 483
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGCCAC 120

Qy        484 TTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAATTG 543
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAATTG 180

Qy        544 TTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGATTCA 603
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGATTCA 240

Qy        604 TCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAG 663
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAG 300

Qy        664 CTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTG 723
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTG 360

Qy        724 AGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTG 783
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTG 420

Qy        784 TTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCAT 843
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCAT 480

Qy        844 TTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCA 882
              |||||||||||||||||||||||||||||||||||||||
Db        481 TTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCA 519
```

```
RESULT 10
BX283437
LOCUS       BX283437                510 bp    mRNA    linear   EST 04-MAR-2003
DEFINITION  BX283437 NIH_MGC_7 Homo sapiens cDNA clone IMAGp958G05802 ;
            IMAGE:3944860, mRNA sequence.
ACCESSION   BX283437
VERSION     BX283437.1  GI:28847891
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 510)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
```

```
            RZPD; IMAGp958G05802.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            pCMV-M13u, Primer sequence: CGTTGTAAAACGACGGCCAGT.
FEATURES             Location/Qualifiers
     source          1. .510
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp958G05802 ; IMAGE:3944860"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_7"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match           37.4%;  Score 496;  DB 13;  Length 510;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-78;
  Matches  496;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 74

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 134

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 194

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 254

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 314

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 374

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 434

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 494

Qy        481 CACTTGTAGCTGGTTT 496
              ||||||||||||||||
Db        495 CACTTGTAGCTGGTTT 510
```

```
RESULT 11
AL138363/c
LOCUS       AL138363                501 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp762F1316_s1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762F1316 3', mRNA sequence.
ACCESSION   AL138363
VERSION     AL138363.1  GI:6855044
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            r1 sequence also available.
            This clone (DKFZp762F1316) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp762F1316"
                     /tissue_type="melanoma (MeWo cell line)"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="762 (synonym: hmel2)"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

  Query Match           36.9%;  Score 489;  DB 9;  Length 501;
  Best Local Similarity  100.0%;  Pred. No. 3.7e-77;
  Matches  489;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        820 GCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCC 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCC 442

Qy        880 CCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCC 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 CCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCC 382

Qy        940 ACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAG 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAG 322

Qy       1000 TGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGG 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 TGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGG 262

Qy       1060 GCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAG 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAG 202

Qy       1120 CTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAA 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAA 142

Qy       1180 GGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTAT 1239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        141 GGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTAT 82

Qy       1240 TTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTA 1299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTA 22

Qy       1300 CAGATGTCA 1308
              |||||||||
Db         21 CAGATGTCA 13
```

RESULT 12
AW515028/c

```
LOCUS       AW515028                470 bp    mRNA    linear   EST 03-MAR-2000
DEFINITION  xu90e05.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2808992 3',
            mRNA sequence.
ACCESSION   AW515028
VERSION     AW515028.1  GI:7153110
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40UP from Gibco
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2808992"
                     /tissue_type="moderately-differentiated endometrial
                     adenocarcinoma, 3 pooled tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Ut2"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.85 kb. Life Technologies catalog #:
                     11539-012"
ORIGIN
```

```
  Query Match             34.3%;  Score 455;  DB 10;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-71;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        849 TAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCA 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCA 396

Qy        909 ATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTC 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 ATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTC 336

Qy        969 TATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTC 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 TATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTC 276

Qy       1029 CCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGAC 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 CCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGAC 216

Qy       1089 ACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 ACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTG 156

Qy       1149 CCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCA 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCA 96

Qy       1209 ACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGA 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 ACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGA 36

Qy       1269 GTTATGATCAATTTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||||||||||||||
Db         35 GTTATGATCAATTTGTATTAAATTAAAGTTACAGA 1
```

RESULT 13
AI457621/c

```
LOCUS       AI457621                486 bp    mRNA    linear   EST 13-APR-1999
DEFINITION  tj63d11.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2146197 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AI457621
VERSION     AI457621.1  GI:4310490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 748   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 453.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2146197"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
ORIGIN
```

```
  Query Match             33.2%;  Score 440;  DB 9;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-68;
```

```
  Matches  440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        864 GTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCG 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCG 382

Qy        924 GGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAG 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAG 322

Qy        984 GTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTT 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTT 262

Qy       1044 TCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCC 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCC 202

Qy       1104 AAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGG 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGG 142

Qy       1164 CCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGG 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 CCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGG 82

Qy       1224 GCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTG 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 GCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTG 22

Qy       1284 TATTAAATTAAAGTTACAGA 1303
              ||||||||||||||||||||
Db          21 TATTAAATTAAAGTTACAGA 2
```

```
RESULT 14
CA434008/c
LOCUS       CA434008               451 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-DF0-arq-p-13-0-UI.s1 NCI_CGAP_DF0 Homo sapiens cDNA clone
            UI-H-DF0-arq-p-13-0-UI 3', mRNA sequence.
ACCESSION   CA434008
VERSION     CA434008.1  GI:24798428
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-DF0-arq-p-13-0-UI"
                     /tissue_type="Subchondral Bone"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DF0"
                     /note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DF0 is a cDNA library containing the following
                     tissue(s): Subchondral Bone. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GTTAAGCGTC.
                     TAG_TISSUE=subchondral bone
                     TAG_LIB=UI-H-DF0
                     TAG_SEQ=GTTAAGCGTC"
ORIGIN
```

```
  Query Match             32.7%;  Score 434;  DB 14;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-67;
  Matches  434;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        870 GGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAG 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 GGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAG 392

Qy        930 GGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGA 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGA 332

Qy        990 ACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTC 1049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 ACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTC 272

Qy       1050 TAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTA 1109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 TAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTA 212

Qy       1110 GCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCC 1169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCC 152

Qy       1170 TTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACC 1229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 TTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACC 92

Qy       1230 CCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAA 1289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 CCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAA 32

Qy       1290 ATTAAAGTTACAGA 1303
              ||||||||||||||
Db         31 ATTAAAGTTACAGA 18
```

```
RESULT 15
BE671191/c
LOCUS       BE671191               477 bp    mRNA    linear   EST 08-SEP-2000
DEFINITION  7e55b02.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3286347 3',
            mRNA sequence.
ACCESSION   BE671191
VERSION     BE671191.1  GI:10031732
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
           Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
           Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL, send email to:
           info@image.llnl.gov
           Seq primer: -40UP from Gibco
           High quality sequence stop: 463.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3286347"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match             32.5%;  Score 431;  DB 10;  Length 477;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-67;
  Matches  431;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        626 AGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCT 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 AGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCT 372

Qy        686 CCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTC 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 CCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTC 312

Qy        746 ACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGAT 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 ACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGAT 252

Qy        806 CCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGT 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGT 192

Qy        866 GGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGG 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGG 132

Qy        926 AGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGT 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 AGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGT 72

Qy        986 GGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTC 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTC 12

Qy       1046 TTTCTAGTCCT 1056
              ||||||||||||
Db         11 TTTCTAGTCCT 1


RESULT 16
AU154016/c
LOCUS       AU154016                557 bp    mRNA    linear   EST 05-AUG-2002
DEFINITION  AU154016 NT2RP3 Homo sapiens cDNA clone NT2RP3004617 3', mRNA
            sequence.
ACCESSION   AU154016
VERSION     AU154016.1  GI:11015537
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T.,
            Sugano,S. and Isogai,T.).
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
            Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M.,
            Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3004617"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /clone_lib="NT2RP3"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
ORIGIN

  Query Match             31.9%;  Score 423;  DB 9;  Length 557;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-65;
  Matches  423;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        885 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 364

Qy        945 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 304

Qy       1005 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 244

Qy       1065 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 184

Qy       1125 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 124

Qy       1185 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 64

Qy       1245 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 4

Qy       1305 GTC 1307
              |||
Db          3 GTC 1
```

```
RESULT 17
AI066431/c
LOCUS       AI066431                423 bp    mRNA    linear   EST 10-NOV-1998
DEFINITION  oq98e10.x1 NCI_CGAP_Co12 Homo sapiens cDNA clone IMAGE:1594410 3',
            mRNA sequence.
ACCESSION   AI066431
VERSION     AI066431.1  GI:3367133
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1019   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 374.
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1594410"
                     /sex="mixed"
                     /tissue_type="colon tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="NCI_CGAP_Co12"
                     /note="Organ: colon; Vector: Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3' Average insert size:  1.2 kb."
ORIGIN

  Query Match           31.7%;  Score 421;  DB 9;  Length 423;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-65;
  Matches  421;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        888 GTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAG 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAG 364

Qy        948 GGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGC 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGC 304

Qy       1008 AAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGAT 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 AAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGAT 244

Qy       1068 TCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACAC 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACAC 184

Qy       1128 AGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGT 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGT 124

Qy       1188 CAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTT 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTT 64

Qy       1248 AAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 AAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 4

Qy       1308 A 1308
              |
Db          3 A 3


RESULT 18
BQ187235
LOCUS       BQ187235                501 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-E-EJ1-ajz-i-06-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
            UI-E-EJ1-ajz-i-06-0-UI 5', mRNA sequence.
ACCESSION   BQ187235
VERSION     BQ187235.1  GI:20362786
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 470-496, >POLY_A#Simple_repeat
            Seq primer: M13 REVERSE.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajz-i-06-0-UI"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EJ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes,
                     AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
```

AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

```
  Query Match             31.7%;  Score 421;  DB 13;  Length 501;
  Best Local Similarity   99.8%;  Pred. No. 3e-65;
  Matches  471;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        832 GGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTT 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTT 60

Qy        892 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 120

Qy        952 GTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGA 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGA 180

Qy       1012 CTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTG 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTG 240

Qy       1072 CACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGG 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGG 300

Qy       1132 GGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGC 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGC 360

Qy       1192 CCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAA 1251
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        361 CCTACCCCAACCCTTCAACTACCAAAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAA 420

Qy       1252 TGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGA 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGA 472
```

RESULT 19
BE795637
LOCUS BE795637 983 bp mRNA linear EST 20-SEP-2000
DEFINITION 601590620F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3944860 5', mRNA sequence.
ACCESSION BE795637
VERSION BE795637.1 GI:10216835
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 983)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM802 row: g column: 05
High quality sequence stop: 791.
FEATURES Location/Qualifiers
source 1. .983
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3944860"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

```
  Query Match             31.3%;  Score 415;  DB 10;  Length 983;
  Best Local Similarity   99.8%;  Pred. No. 1.8e-64;
  Matches  535;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          9 GGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGCCACCCCCA 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGCCACCCCCA 60

Qy         69 CACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGT 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGT 120

Qy        129 GGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTC 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTC 180

Qy        189 CTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAG 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAG 240

Qy        249 CCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAAC 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAAC 300

Qy        309 AAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCAGGAGGCTT 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCAGGAGGCTT 360

Qy        369 CCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCTAGACCCCT 428
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGAAGCAGTTG-TGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCTAGACCCCT 419

Qy        429 TGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGCCACTTGTA 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 TGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGCCACTTGTA 479

Qy        489 GCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAATTGT 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAATTGT 535
```

RESULT 20
BE856182/c
LOCUS BE856182 456 bp mRNA linear EST 29-SEP-2000
DEFINITION 7f89g01.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3304176 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION BE856182
VERSION BE856182.1 GI:10368957
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 456)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

```
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3304176"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr28"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           30.5%;  Score 405;  DB 10;  Length 456;
  Best Local Similarity   99.8%;  Pred. No. 2.1e-62;
  Matches  455;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        847 CCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 CCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATC 397

Qy        907 CAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGC 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 CAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGC 337

Qy        967 TCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 TCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTC 277

Qy       1027 TCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTG 1086
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        276 TCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTATG 217

Qy       1087 ACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGC 157

Qy       1147 TGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 TGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTT 97

Qy       1207 CAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTAT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 CAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTAT 37

Qy       1267 GAGTTATGATCAATTTGTATTAAATTAAAGTTACAG 1302
              ||||||||||||||||||||||||||||||||||||
Db          36 GAGTTATGATCAATTTGTATTAAATTAAAGTTACAG 1

RESULT 21
AA700789/c
LOCUS       AA700789                 456 bp    mRNA    linear   EST 19-DEC-1997
DEFINITION  zi66c04.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:435750 3', mRNA sequence.
ACCESSION   AA700789
VERSION     AA700789.1  GI:2703954
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 439.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:1335519"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:435750"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match           30.4%;  Score 403;  DB 9;  Length 456;
  Best Local Similarity   99.8%;  Pred. No. 4.8e-62;
  Matches  453;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        855 GAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 GAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTG 397

Qy        915 AAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 AAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCC 337

Qy        975 TGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGC 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 TGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGC 277

Qy       1035 CTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 CTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACA 217

Qy       1095 CACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCG 1154
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        216 CACCATCCCAAAGTAGCCGGAAGAGCTAAACATAGGGGGTTCTTAAAATGGCTGCCCCCG 157
```

```
Qy       1155 CCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACC 1214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACC 97

Qy       1215 AGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATG 1274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          96 AGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATG 37

Qy       1275 ATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
             |||||||||||||||||||||||||||||||||
Db          36 ATCAATTTGTATTAAATTAAAGTTACAGATGTCA 3


RESULT 22
CK300567/c
LOCUS       CK300567                502 bp    mRNA    linear   EST 15-DEC-2003
DEFINITION  UI-E-EJ1-ajz-i-06-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
            UI-E-EJ1-ajz-i-06-0-UI 3', mRNA sequence.
ACCESSION   CK300567
VERSION     CK300567.1  GI:39890062
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Greg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/eye.html
             The following repetitive elements were found in this cDNA
            sequence: 1-40, >AT_rich#Low_complexity (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .502
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajz-i-06-0-UI"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EJ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes,
                     AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
                     AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
                     Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
                     library was created for the program, Gene Discovery in the
                     Visual System, supported by National Eye Institute (NEI).
                     TAG_TISSUE=RPE and Choroid
                     TAG_LIB=UI-E-EJ1
                     TAG_SEQ=ACCTA"
ORIGIN

  Query Match             30.1%;  Score 399;  DB 14;  Length 502;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-61;
  Matches  399;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        832 GGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTT 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 GGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTT 434

Qy        892 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 374

Qy        952 GTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGA 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGA 314

Qy       1012 CTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTG 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTG 254

Qy       1072 CACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGG 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 CACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGG 194

Qy       1132 GGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGC 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGC 134

Qy       1192 CCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCC 1230
              |||||||||||||||||||||||||||||||||||||||
Db        133 CCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCC 95


RESULT 23
AI681374/c
LOCUS       AI681374                512 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tx46a12.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2272606 3',
            mRNA sequence.
ACCESSION   AI681374
VERSION     AI681374.1  GI:4891556
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 512)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1330   Std Error: 0.00
            Seq primer: -40UP from Gibco
```

```
          High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .512
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2272606"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           29.5%;  Score 391;  DB 9;  Length 512;
  Best Local Similarity   99.8%;  Pred. No. 5.4e-60;
  Matches  511;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        798 AGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAA 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 AGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAA 453

Qy        858 GTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAG 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 GTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAG 393

Qy        918 AACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGC 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 AACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGC 333

Qy        978 CCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGG-CCT 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        332 CCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCCT 273

Qy       1037 TGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACA 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 TGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACA 213

Qy       1097 CCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCC 1156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         212 CCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCC 153

Qy       1157 ACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAG 1216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         152 ACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAG 93

Qy       1217 AATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGAT 1276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          92 AATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGAT 33

Qy       1277 CAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||
Db          32 CAATTTGTATTAAATTAAAGTTACAGATGTCA 1


RESULT 24
R71654/c
LOCUS       R71654                 431 bp    mRNA    linear   EST 01-JUN-1995
DEFINITION  yi53g05.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:143000 3', mRNA sequence.
ACCESSION   R71654
VERSION     R71654.1  GI:845171
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 855
            High quality sequence stops: 325
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 855   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:552113"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:143000"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "
ORIGIN

  Query Match           28.5%;  Score 378;  DB 14;  Length 431;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-57;
  Matches  378;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        918 AACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGC 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGC 332

Qy        978 CCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTT 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTT 272

Qy       1038 GACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACAC 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 GACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACAC 212

Qy       1098 CATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCA 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         211 CATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCA 152

Qy       1158 CCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGA 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         151 CCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGA 92

Qy       1218 ATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          91 ATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATC 32
```

```
Qy       1278 AATTTGTATTAAATTAAA 1295
              ||||||||||||||||||
Db         31 AATTTGTATTAAATTAAA 14


RESULT 25
BE244684
LOCUS       BE244684                497 bp    mRNA    linear   EST 03-OCT-2001
DEFINITION  TCBAP2E0532 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP0532, mRNA
            sequence.
ACCESSION   BE244684
VERSION     BE244684.1  GI:9096426
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R.,
            Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Citation: Carninci,P. and Hayashizaki,Y. High efficiency
            full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TCBAP0532"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /lab_host="DH10B"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project=TCBA"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and XhoI
                     and directionally cloned into the BamH1 and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
ORIGIN

  Query Match             27.9%;  Score 370;  DB 10;  Length 497;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-56;
  Matches  370;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        589 CCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGC 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGC 187

Qy        649 CACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAG 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        188 CACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAG 247

Qy        709 CCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTT 307

Qy        769 TTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAG 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 TTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAG 367

Qy        829 TCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCG 427

Qy        889 TTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGG 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 TTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGG 487

Qy        949 GTTGTCCTGC 958
              ||||||||||
Db        488 GTTGTCCTGC 497


RESULT 26
CA434086/c
LOCUS       CA434086                424 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-DF0-arq-p-13-0-UI.s2 NCI_CGAP_DF0 Homo sapiens cDNA clone
            UI-H-DF0-arq-p-13-0-UI 3', mRNA sequence.
ACCESSION   CA434086
VERSION     CA434086.1  GI:24798506
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .424
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-DF0-arq-p-13-0-UI"
                     /tissue_type="Subchondral Bone"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DF0"
                     /note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DF0 is a cDNA library containing the following
                     tissue(s): Subchondral Bone. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GTTAAGCGTC.
```

```
                    TAG_TISSUE=subchondral bone
                    TAG_LIB=UI-H-DF0
                    TAG_SEQ=GTTAAGCGTC"
ORIGIN

  Query Match             24.0%;  Score 319;  DB 14;  Length 424;
  Best Local Similarity   99.7%;  Pred. No. 2.7e-47;
  Matches  369;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        934 GAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTG 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTG 328

Qy        994 GAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGT 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGT 268

Qy       1054 CCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCG 1113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCG 208

Qy       1114 GAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGG 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGG 148

Qy       1174 GCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAG 1233
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        147 GCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATTTGGGCCACCCCAG 88

Qy       1234 CAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTA 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          87 CAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTA 28

Qy       1294 AAGTTACAGA 1303
              ||||||||||
Db          27 AAGTTACAGA 18


RESULT 27
R71157
LOCUS       R71157                   543 bp    mRNA    linear   EST 01-JUN-1995
DEFINITION  yi53g05.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:143000 5', mRNA sequence.
ACCESSION   R71157
VERSION     R71157.1  GI:844674
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 855
            High quality sequence stops: 311
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 855   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 311.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:552113"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:143000"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares placenta Nb2HP"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "
ORIGIN

  Query Match             23.0%;  Score 305;  DB 14;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-45;
  Matches  305;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        560 AGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGAC 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGAC 60

Qy        620 CAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAAC 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAAC 120

Qy        680 TAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCC 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCC 180

Qy        740 CACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAG 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAG 240

Qy        800 TCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGT 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGT 300

Qy        860 TAGGG 864
              |||||
Db        301 TAGGG 305


RESULT 28
AI383065/c
LOCUS       AI383065                 304 bp    mRNA    linear   EST 18-MAR-1999
DEFINITION  tc20g06.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2064442
            3', mRNA sequence.
ACCESSION   AI383065
VERSION     AI383065.1  GI:4195846
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 304)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 503   Std Error: 0.00
            Seq primer: -40UP from Gibco
```

```
          High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .304
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2064442"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
ORIGIN

  Query Match             22.9%;  Score 304;  DB 9;  Length 304;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-44;
  Matches  304;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        997 GAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCT 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCT 245

Qy       1057 GGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAA 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAA 185

Qy       1117 GAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCA 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         184 GAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCA 125

Qy       1177 AAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         124 AAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAG 65

Qy       1237 TATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAG 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          64 TATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAG 5

Qy       1297 TTAC 1300
              ||||
Db           4 TTAC 1


RESULT 29
BE245533/c
LOCUS       BE245533                376 bp    mRNA    linear   EST 03-OCT-2001
DEFINITION  TCBAP1D3276 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP3276, mRNA
            sequence.
ACCESSION   BE245533
VERSION     BE245533.1  GI:9097279
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R.,
            Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Citation: Carninci,P. and Hayashizaki,Y. High efficiency
            full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .376
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TCBAP3276"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /lab_host="DH10B"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project=TCBA"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and XhoI
                     and directionally cloned into the BamH1 and Sa1I sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
ORIGIN

  Query Match             22.5%;  Score 299;  DB 10;  Length 376;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-44;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        927 GAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTG 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTG 317

Qy        987 GGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCT 1046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCT 257

Qy       1047 TTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAA 1106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAA 197

Qy       1107 GTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCT 1166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCT 137

Qy       1167 CCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGC 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 CCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGC 78


RESULT 30
AA644653
LOCUS       AA644653                385 bp    mRNA    linear   EST 02-MAR-1998
DEFINITION  af74f11.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1047789
            5', mRNA sequence.
ACCESSION   AA644653
VERSION     AA644653.1  GI:2569871
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 499   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .385
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1047789"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
ORIGIN

  Query Match           21.6%;  Score 286;  DB 9;  Length 385;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-41;
  Matches  286;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        967 TCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 TCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTC 159

Qy       1027 TCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTG 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 TCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTG 219

Qy       1087 ACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         220 ACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGC 279

Qy       1147 TGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         280 TGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTT 339

Qy       1207 CAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAAT 1252
              |||||||||||||||||||||||||||||||||||||||||||||
Db         340 CAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAAT 385


RESULT 31
AW139822/c
LOCUS       AW139822                434 bp    mRNA    linear   EST 30-OCT-1999
DEFINITION  UI-H-BI1-adw-c-09-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
            IMAGE:2718328 3', mRNA sequence.
ACCESSION   AW139822
VERSION     AW139822.1  GI:6144540
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 434)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2718328"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub3"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NCI_CGAP_Sub3 library is a subtracted library derived from
                     the NCI_CGAP_Sub1 library, which is a subtracted library
                     derived from BI. BI constitutes a mixture of 21
                     normalized or subtracted NCI_CGAP libraries:
                     NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
                     NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
                     NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
                     NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2,
                     NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
                     NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
                     NCI_CGAP_Brn25. These 21 libraries were pooled and a
                     single-stranded DNA preparation of the resulting mixture
                     was used as a tracer in a subtractive hybridization with
                     a driver whose composition is detailed below:
                     NCI_CGAP_Kid3 pool 1    LLAM 3334-3337, 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911,
                     1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1
                     LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs
                     1323912-1325831, 1471368-1472903, 1492104-1493255);
                     NCI_CGAP_Lu5 pool 1     LLAM 3575-3582, 3851-3854 (IMAGE
                     CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4
                     pool 1     LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
                     CloneIDs 1257096-1258631, 1469064-1470983,
                     1475592-1476743); NCI_CGAP_Pr22 pool 1    LLAM 2457-2459,
                     2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
                     1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
                     LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,
                     1144584-1145351). Subtraction was performed as previously
                     described [Bonaldo, Lennon & Soares (1996): Normalization
                     and Subtraction: Two Approaches To Facilitate Gene
                     Discovery. Genome Research 6, 791-806.
                     TAG_TISSUE=brain
                     TAG_LIB=NCI_CGAP_Brn23
                     TAG_SEQ=ATATC"
ORIGIN

  Query Match           21.5%;  Score 285;  DB 10;  Length 434;
  Best Local Similarity  99.7%;  Pred. No. 2.4e-41;
  Matches  335;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        973 CCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCG 1032
```

```
Db        355 CCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCG 296

Qy       1033 GCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAA 1092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAA 236

Qy       1093 CACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCC 1152
             ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        235 CACACCATCCCAAAGTAGTCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCC 176

Qy       1153 CGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTA 1212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTA 116

Qy       1213 CCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTA 1272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTA 56

Qy       1273 TGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
             ||||||||||||||||||||||||||||||||||
Db          55 TGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 20


RESULT 32
H47130
LOCUS       H47130                  431 bp    mRNA    linear   EST 16-AUG-1995
DEFINITION  yp73a07.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:193044 5', mRNA sequence.
ACCESSION   H47130
VERSION     H47130.1  GI:923182
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 935
            High quality sequence stops: 322
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 935   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 322.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3762065"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:193044"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             21.3%;  Score 283;  DB 14;  Length 431;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-41;
  Matches  283;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        466 GCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCC 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCC 60

Qy        526 TGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAA 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAA 120

Qy        586 CTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGA 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGA 180

Qy        646 AGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGC 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGC 240

Qy        706 CAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACT 748
              |||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACT 283


RESULT 33
H78426
LOCUS       H78426                  440 bp    mRNA    linear   EST 09-NOV-1995
DEFINITION  yu22c04.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:234534 5', mRNA sequence.
ACCESSION   H78426
VERSION     H78426.1  GI:1056515
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 929
            High quality sequence stops: 356
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 929   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 356.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3787711"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:234534"
```

/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 21.3%; Score 283; DB 14; Length 440;
Best Local Similarity 100.0%; Pred. No. 5.2e-41;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        466 GCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCC 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCC 71

Qy        526 TGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAA 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAA 131

Qy        586 CTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGA 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 CTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGA 191

Qy        646 AGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGC 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 AGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGC 251

Qy        706 CAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACT 748
              |||||||||||||||||||||||||||||||||||||||||||
Db        252 CAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACT 294
```

RESULT 34
AA037876
LOCUS AA037876 513 bp mRNA linear EST 01-FEB-1997
DEFINITION zf04f10.r1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:375979 5', mRNA sequence.
ACCESSION AA037876
VERSION AA037876.1 GI:1513012
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 513)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 768 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 408.
FEATURES Location/Qualifiers
source 1. .513
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1284235"
/db_xref="taxon:9606"
/clone="IMAGE:375979"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_heart_NbHH19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

ORIGIN

Query Match 21.1%; Score 280; DB 9; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.5e-40;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        622 GGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTA 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTA 64

Qy        682 GGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCA 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCA 124

Qy        742 CTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTC 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTC 184

Qy        802 AGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTA 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 AGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTA 244

Qy        862 GGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACA 901
              ||||||||||||||||||||||||||||||||||||||||
Db        245 GGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACA 284
```

RESULT 35
AW070327/c
LOCUS AW070327 543 bp mRNA linear EST 20-OCT-2000
DEFINITION xa10c02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2567906 3', mRNA sequence.
ACCESSION AW070327
VERSION AW070327.1 GI:6025325
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

```
           Insert Length: 850   Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 464.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2567906"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match            20.9%;  Score 277;  DB 9;  Length 543;
  Best Local Similarity   99.1%;  Pred. No. 4.8e-40;
  Matches  527;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        772 CAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCT 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCT 484

Qy        832 GGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTT 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTT 424

Qy        892 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 951
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TTTTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 364

Qy        952 GTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGA 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGA 304

Qy       1012 CTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGGCCTAGATTCTG 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        303 CTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGGCCTAGATTNTG 244

Qy       1072 CACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGG 1131
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CACTTGGGGTCTTTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGG 184

Qy       1132 GGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGC 1191
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GGTTTTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGC 124

Qy       1192 CCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAA 1251
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        123 CCTACCCCAACCCTTCAACTACCAGAATTTGGGCCACCCCAGCAGTATTTTTATTTAAAA 64

Qy       1252 TGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGA 12


RESULT 36
T85444
LOCUS       T85444                 389 bp    mRNA    linear   EST 17-MAR-1995
DEFINITION  yd82c03.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:114724 5', mRNA sequence.
ACCESSION   T85444
VERSION     T85444.1  GI:713796
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 948
            High quality sequence stops: 261 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 948   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 261.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:470341"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:114724"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match            20.3%;  Score 269;  DB 14;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-38;
  Matches  269;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        467 CAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCT 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCT 60

Qy        527 GGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAAC 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAAC 120

Qy        587 TCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAA 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAA 180

Qy        647 GCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCC 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCC 240

Qy        707 AGCCAATAGATGTCCTGAGGCCTGCCCCT 735
              |||||||||||||||||||||||||||||
Db        241 AGCCAATAGATGTCCTGAGGCCTGCCCCT 269
```

RESULT 37
BE219390/c
LOCUS BE219390 470 bp mRNA linear EST 03-JUL-2000
DEFINITION hv57c03.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3177508 3', mRNA sequence.
ACCESSION BE219390
VERSION BE219390.1 GI:8906708
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 470)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 446.
FEATURES Location/Qualifiers
source 1. .470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3177508"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Query Match 20.3%; Score 269; DB 10; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       988 GAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTT 1047
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       318 GAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTT 259

Qy      1048 TCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAG 1107
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       258 TCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAG 199

Qy      1108 TAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTC 1167
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       198 TAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTC 139

Qy      1168 CCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCA 1227
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       138 CCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCA 79

Qy      1228 CCCCAGCAGTATTTTTATTTAAAATGTTG 1256
            |||||||||||||||||||||||||||||
Db        78 CCCCAGCAGTATTTTTATTTAAAATGTTG 50
```

RESULT 38
BE245721/c
LOCUS BE245721 337 bp mRNA linear EST 03-OCT-2001
DEFINITION TCBAP1D1925 Pediatric pre-B cell acute lymphoblastic leukemia Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP1925, mRNA sequence.
ACCESSION BE245721
VERSION BE245721.1 GI:9097468
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 337)
AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
TITLE Pediatric Leukemia cDNA Sequencing Project
JOURNAL Unpublished (2000)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Citation: Carninci,P. and Hayashizaki,Y. High efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Seq primer: M13 primer.
FEATURES Location/Qualifiers
source 1. .337
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCBAP1925"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/clone_lib="Pediatric pre-B cell acute lymphoblastic leukemia Baylor-HGSC project=TCBA"
/note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN 3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand was primed with a BamH1-dC primer [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3']. Double-stranded cDNA was then digested with BamH1 and XhoI and directionally cloned into the BamH1 and Sal1 sites of lambda pSB vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T, Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y, High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)"
ORIGIN

Query Match 19.6%; Score 260; DB 10; Length 337;
Best Local Similarity 100.0%; Pred. No. 7.3e-37;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       967 TCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTC 1026
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       336 TCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTC 277

Qy      1027 TCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTG 1086
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       276 TCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTG 217

Qy      1087 ACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGC 1146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        216 ACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGC 157

Qy       1147 TGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 TGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTT 97

Qy       1207 CAACTACCAGAATCTGGGCC 1226
              ||||||||||||||||||||
Db          96 CAACTACCAGAATCTGGGCC 77


RESULT 39
BF696473
LOCUS       BF696473                749 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602125145F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4282091 5',
            mRNA sequence.
ACCESSION   BF696473
VERSION     BF696473.1  GI:11981881
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1111  row: j  column: 12
            High quality sequence stop: 560.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4282091"
                     /tissue_type="primitive neuroectoderm"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_56"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.65 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

  Query Match             19.2%;  Score 255;  DB 10;  Length 749;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-36;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         70 ACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGTG 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGTG 204

Qy        130 GCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTCC 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTCC 264

Qy        190 TGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAGC 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAGC 324

Qy        250 CTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAACA 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 CTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAACA 384

Qy        310 AGAGTCCAGCCAACC 324
              ||||||||||||||||
Db        385 AGAGTCCAGCCAACC 399


RESULT 40
BU754096/c
LOCUS       BU754096                270 bp    mRNA    linear   EST 10-OCT-2002
DEFINITION  UI-1-BB1p-atn-h-07-0-UI.s1 NCI_CGAP_P16 Homo sapiens cDNA clone
            UI-1-BB1p-atn-h-07-0-UI 3', mRNA sequence.
ACCESSION   BU754096
VERSION     BU754096.1  GI:23712667
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .270
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BB1p-atn-h-07-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="Full Term"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_P16"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_P16 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human full term
                     TAG_LIB=UI-1-BB1p
                     TAG_SEQ=AGGAA"
ORIGIN

  Query Match             18.6%;  Score 247;  DB 13;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-34;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1057 GGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAA 1116
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAA 203

Qy       1117 GAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCA 1176
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCA 143

Qy       1177 AAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAG 1236
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAG 83

Qy       1237 TATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAG 1296
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 TATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAG 23

Qy       1297 TTACAGA 1303
                |||||||
Db          22 TTACAGA 16


RESULT 41
AW964518
LOCUS       AW964518                 378 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST376591 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW964518
VERSION     AW964518.1  GI:8154354
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
            Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 202
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .378
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGH"
                     /note="Vector: pBluescriptSKm"
ORIGIN

  Query Match            18.2%;  Score 241;  DB 10;  Length 378;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-33;
  Matches  241;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        640 GCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGT 699
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGT 60

Qy        700 CCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAA 759
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAA 120

Qy        760 TCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCT 819
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCT 180

Qy        820 GCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCC 879
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCC 240

Qy        880 C 880
                |
Db        241 C 241


RESULT 42
AA191674/c
LOCUS       AA191674                 350 bp    mRNA    linear   EST 12-MAR-1998
DEFINITION  zq43g04.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:632502 3', mRNA sequence.
ACCESSION   AA191674
VERSION     AA191674.1  GI:1780345
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 3281   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham.
FEATURES             Location/Qualifiers
     source          1. .350
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:5182585"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:632502"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match            17.0%;  Score 226;  DB 9;  Length 350;
  Best Local Similarity  99.6%;  Pred. No. 6.3e-31;
  Matches  276;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1031 CGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACAC 1090
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTNTGACAC 218

Qy       1091 AACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCC 1150
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCC 158

Qy       1151 CCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAAC 1210
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAAC 98

Qy       1211 TACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGT 1270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 TACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGT 38
```

```
Qy       1271 TATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
              |||||||||||||||||||||||||||||||||||||
Db         37 TATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1
```

```
RESULT 43
BE817016/c
LOCUS       BE817016                334 bp    mRNA    linear   EST 21-SEP-2000
DEFINITION  RC0-BN0248-310700-025-f05 BN0248 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE817016
VERSION     BE817016.1  GI:10249250
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 334)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC0-BN0248-310
            700-025-f05&t3=2000-07-31&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 26
            High quality sequence stop: 334.
FEATURES             Location/Qualifiers
     source          1. .334
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="BN0248"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN
```

```
  Query Match             17.0%;  Score 225;  DB 10;  Length 334;
  Best Local Similarity   100.0%;  Pred. No. 9.9e-31;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          83 TGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGTGGCCTGTGGAAGGC 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 TGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGATGATGCTGTGGCCTGTGGAAGGC 263

Qy         143 ACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCG 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 ACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCG 203

Qy         203 AGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTG 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 AGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTG 143

Qy         263 CAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAA 307
               |||||||||||||||||||||||||||||||||||||||||||||
Db         142 CAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAGCTCTGAA 98
```

```
RESULT 44
BG912772
LOCUS       BG912772                405 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602807993F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4940254
            5', mRNA sequence.
ACCESSION   BG912772
VERSION     BG912772.1  GI:14293248
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10878  row: i  column: 23
            High quality sequence stop: 405.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4940254"
                     /tissue_type="anaplastic oligodendroglioma with 1p/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Brn67"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
```

```
  Query Match             16.7%;  Score 221;  DB 12;  Length 405;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-30;
  Matches  271;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1037 TGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACA 1096
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         116 TGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACA 175

Qy        1097 CCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCC 1156
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         176 CCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCC 235

Qy        1157 ACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAG 1216
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         236 ACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAG 295

Qy        1217 AATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGAT 1276
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         296 AATCTGGGCCACCCCAGCAGTATTTTTATTTACAATGTTGCCCATTTTATGAGTTATGAT 355

Qy        1277 CAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
               ||||||||||||||||||||||||||||||||
```

```
Db        356 CAATTTGTATTAAATTAAAGTTACAGATGTCA 387


RESULT 45
AL554421
LOCUS       AL554421                768 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL554421 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI084YE24 5-PRIME, mRNA sequence.
ACCESSION   AL554421
VERSION     AL554421.2  GI:31276232
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12895186.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 221.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI084BC12QP1&cluster=221.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI084BC12QP1.
FEATURES             Location/Qualifiers
     source          1. .768
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI084YE24"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             16.7%;  Score 221;  DB 9;  Length 768;
  Best Local Similarity   98.9%;  Pred. No. 2.2e-30;
  Matches  471;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        831 TGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTT 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 TGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTT 328

Qy        891 TTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGT 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGT 388

Qy        951 TGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAG 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAG 448

Qy       1011 ACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCT 1070
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 AMTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCT 508

Qy       1071 GCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGG 1130
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        509 GCACTTGGGGTCTCTGACACAACAAACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGG 568

Qy       1131 GGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAG 1190
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        569 GGGTTCTTAAAATGGCTGCACCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAG 628

Qy       1191 CCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAA 1250
              | ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        629 CACTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTWTTTTTATTTAAA 688

Qy       1251 ATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGT 1306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 ATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGT 744


RESULT 46
BU527114
LOCUS       BU527114                931 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT_10155963 NIH_MGC_101 Homo sapiens cDNA clone
            IMAGE:6536868 5', mRNA sequence.
ACCESSION   BU527114
VERSION     BU527114.1  GI:22837555
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2698  row: o  column: 12
            High quality sequence stop: 552.
FEATURES             Location/Qualifiers
     source          1. .931
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6536868"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_101"
                     /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             16.0%;  Score 212;  DB 13;  Length 931;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-29;
  Matches  212;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 460

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 520

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 580

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGG 212
```

```
                    ||||||||||||||||||||||||||||||
Db        581 CCACCTTCCTGCCCACAGGCCGAGGGACAGGG 612


RESULT 47
AI864353/c
LOCUS       AI864353                310 bp    mRNA    linear   EST 30-AUG-1999
DEFINITION  wg78f06.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2371235 3', mRNA sequence.
ACCESSION   AI864353
VERSION     AI864353.1  GI:5528460
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2371235"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
ORIGIN

  Query Match             15.9%;  Score 211;  DB 9;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 3e-28;
  Matches  211;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1093 CACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCC 1152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 CACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCC 153

Qy       1153 CGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTA 1212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         152 CGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTA 93

Qy       1213 CCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTA 1272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          92 CCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTA 33

Qy       1273 TGATCAATTTGTATTAAATTAAAGTTACAGA 1303
             ||||||||||||||||||||||||||||||
Db          32 TGATCAATTTGTATTAAATTAAAGTTACAGA 2


RESULT 48
BQ025918/c
LOCUS       BQ025918                270 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-1-BB1p-ayg-d-07-0-UI.s1 NCI_CGAP_Pl6 Homo sapiens cDNA clone
            UI-1-BB1p-ayg-d-07-0-UI 3', mRNA sequence.
ACCESSION   BQ025918
VERSION     BQ025918.1  GI:19761197
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .270
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BB1p-ayg-d-07-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="Full Term"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Pl6"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Pl6 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human full term
                     TAG_LIB=UI-1-BB1p
                     TAG_SEQ=AGGAA"
ORIGIN

  Query Match             14.8%;  Score 196;  DB 12;  Length 270;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-25;
  Matches  246;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1057 GGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAA 1116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAA 203

Qy       1117 GAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCA 1176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCA 143

Qy       1177 AAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAG 1236
             ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        142 AAAGGAATTGTCAGCCCTACCCCAACCCTTCAAATACCAGAATCTGGGCCACCCCAGCAG 83
```

```
Qy       1237 TATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAG 1296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 TATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAG 23

Qy       1297 TTACAGA 1303
             |||||||
Db         22 TTACAGA 16


RESULT 49
AA037877/c
LOCUS       AA037877                488 bp    mRNA    linear   EST 01-FEB-1997
DEFINITION  zf04f10.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:375979 3', mRNA sequence.
ACCESSION   AA037877
VERSION     AA037877.1  GI:1513013
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 768   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 376.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:1284235"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:375979"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'},
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
ORIGIN

  Query Match           14.6%;  Score 194;  DB 9;  Length 488;
  Best Local Similarity  99.6%;  Pred. No. 1.9e-25;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1063 TAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTA 1122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTA 209

Qy       1123 AACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGA 1182
             ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AACACAGGGGGTTNTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGA 149

Qy       1183 ATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTT 1242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTT 89

Qy       1243 TATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAG 1302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          88 TATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAG 29

Qy       1303 ATGTC 1307
             |||||
Db          28 ATGTC 24


RESULT 50
BM678923/c
LOCUS       BM678923                401 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-EO0-ahx-d-01-0-UI.s1 UI-E-EO0 Homo sapiens cDNA clone
            UI-E-EO0-ahx-d-01-0-UI 3', mRNA sequence.
ACCESSION   BM678923
VERSION     BM678923.1  GI:18988819
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EO0-ahx-d-01-0-UI"
                     /tissue_type="fetal eye"
                     /dev_stage="fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EO0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EO0 is a cDNA library containing the following
                     tissue(s): fetal eye. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
```

with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=human fetal eye
TAG_LIB=UI-E-EO0
TAG_SEQ=CGCGTATACC"

ORIGIN

Query Match 13.7%; Score 182; DB 12; Length 401;
Best Local Similarity 99.1%; Pred. No. 2.8e-23;
Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        969 TATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTC 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTC 291

Qy       1029 CCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGAC 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        290 CCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTGTGAC 231

Qy       1089 ACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTG 1148
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        230 ACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTGTTAAAATGGCTG 171

Qy       1149 CCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCA 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 CCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCA 111

Qy       1209 ACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGA 1268
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 AATACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGA 51

Qy       1269 GTTATGATCAATTTGTATTAAATTAAAGTTACAGA 1303
              ||||||||||||||||||||||||||||||||||
Db          50 GTTATGATCAATTTGTATTAAATTAAAGTTACAGA 16
```

Search completed: September 17, 2004, 05:40:47
Job time : 3287 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 00:14:08 ; Search time 5150 Seconds
(without alignments)
11168.195 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtggtcacaggtag.....aaaaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1327	100.0	3259	9	BC001222	BC001222 Homo sapi
2		1327	100.0	3259	9	BC007999	BC007999 Homo sapi
3		1327	100.0	3259	9	BC011689	BC011689 Homo sapi
4		1327	100.0	3259	9	BC012152	BC012152 Homo sapi
5		1327	100.0	3826	6	AX491279	AX491279 Sequence
6		1308	98.6	23433	6	AX491283	AX491283 Sequence
7	c	1308	98.6	164950	2	AC022262	AC022262 Homo sapi
8		1307	98.5	3243	6	AX775831	AX775831 Sequence
9		1257	94.7	30676	6	AX491284	AX491284 Sequence
10	c	1257	94.7	64693	9	AL662907	AL662907 Human DNA
11		1256	94.6	3436	6	AX877131	AX877131 Sequence
12		1256	94.6	3436	6	BD156501	BD156501 Primer fo
13		1256	94.6	3436	9	AK001621	AK001621 Homo sapi
14		1053	79.4	136095	2	AF161326	AF161326 Homo sapi
15		726	54.7	30625	6	AX491281	AX491281 Sequence
16		452	34.1	2207	6	AX775827	AX775827 Sequence
17		452	34.1	2207	6	AX776017	AX776017 Sequence
18		452	34.1	2246	6	AX775829	AX775829 Sequence
19	c	423	31.9	557	6	AX873748	AX873748 Sequence
20	c	423	31.9	557	6	BD153810	BD153810 Primer fo
21		336	25.3	3148	9	AK122896	AK122896 Homo sapi
22	c	62	4.7	250	11	G15084	G15084 human STS S
23		29	2.2	614	6	AX388847	AX388847 Sequence
24	c	29	2.2	1652	6	AX771586	AX771586 Sequence
25	c	29	2.2	162959	2	AC019255	AC019255 Homo sapi
26		29	2.2	178217	9	AC022080	AC022080 Homo sapi
27	c	29	2.2	180338	10	AC090496	AC090496 Mus muscu
28		28	2.1	723	10	BC060953	BC060953 Mus muscu
29		28	2.1	136037	9	AC004104	AC004104 Homo sapi
30		28	2.1	137648	2	AC024316	AC024316 Homo sapi
31	c	28	2.1	158608	9	AC051635	AC051635 Homo sapi
32		28	2.1	187557	2	AC069032	AC069032 Homo sapi
33	c	28	2.1	193978	2	AC068678	AC068678 Homo sapi
34		28	2.1	209462	2	AC011814	AC011814 Homo sapi
35		27	2.0	4150	10	BC060625	BC060625 Mus muscu
36		27	2.0	4685	9	HSM803417	AL832110 Homo sapi
37	c	27	2.0	88643	8	AC022522	AC022522 Arabidops
38		27	2.0	105199	9	AP001576	AP001576 Homo sapi
39		27	2.0	133897	9	AC112138	AC112138 Homo sapi
40	c	27	2.0	152772	9	AC099326	AC099326 Homo sapi
41	c	27	2.0	153448	9	AC063923	AC063923 Homo sapi
42	c	27	2.0	174531	9	AC104449	AC104449 Homo sapi
43		27	2.0	177672	2	AC010959	AC010959 Homo sapi
44		27	2.0	179585	9	AC106783	AC106783 Homo sapi
45		27	2.0	189013	6	AX706971	AX706971 Sequence
46		27	2.0	189013	6	AX707901	AX707901 Sequence
47		27	2.0	189032	9	AC025277	AC025277 Homo sapi
48	c	27	2.0	199045	2	AC110206	AC110206 Mus muscu
49	c	27	2.0	204855	10	AC133100	AC133100 Mus muscu
50		27	2.0	232802	2	AC099108	AC099108 Rattus no
51		26	2.0	651	5	AB033881	AB033881 Coturnix
52		26	2.0	698	8	BOIFABP2	X97677 B.oleracea
53		26	2.0	913	10	BC061227	BC061227 Mus muscu
54		26	2.0	931	10	BC049725	BC049725 Mus muscu
55		26	2.0	1077	9	BC062787	BC062787 Homo sapi
56		26	2.0	1078	8	AY173062	AY173062 Chrysanth
57		26	2.0	1186	10	BC063154	BC063154 Rattus no
58		26	2.0	1238	3	AK115014	AK115014 Ciona int
59		26	2.0	1570	10	BC049624	BC049624 Mus muscu
60		26	2.0	1888	5	BC056828	BC056828 Danio rer
61		26	2.0	2250	10	BC059877	BC059877 Mus muscu
62		26	2.0	2619	10	BC019375	BC019375 Mus muscu
63		26	2.0	2893	10	BC042763	BC042763 Mus muscu
64		26	2.0	2955	10	BC021442	BC021442 Mus muscu
65		26	2.0	3270	9	BC050552	BC050552 Homo sapi

	66	26	2.0	4059	10	BC056981
	67	26	2.0	4798	10	BC056365
	68	26	2.0	49830	6	AX059541
c	69	26	2.0	78239	8	AB012241
c	70	26	2.0	86424	2	AC010222
	71	26	2.0	96685	8	AC007087
	72	26	2.0	115435	9	AC097626
c	73	26	2.0	130965	9	AC026773
	74	26	2.0	142290	2	AC102011
c	75	26	2.0	142290	2	AC102011
c	76	26	2.0	145453	8	AC012477
c	77	26	2.0	148043	9	AC008541
c	78	26	2.0	157848	9	AC079763
c	79	26	2.0	174645	2	AC136857
c	80	26	2.0	183147	8	AC012392
	81	26	2.0	184132	2	AC120345
	82	26	2.0	187162	2	AC139042
	83	26	2.0	192309	2	BX545915
	84	26	2.0	192586	10	AL645972
c	85	26	2.0	195208	10	AL591146
c	86	26	2.0	199662	2	AC131241
	87	26	2.0	199987	8	ATCHRIV15
c	88	26	2.0	201174	10	AC113953
c	89	26	2.0	207136	10	AC122268
	90	26	2.0	211393	2	AC119244
	91	26	2.0	211803	2	AC117813
	92	26	2.0	219288	2	AC131321
c	93	26	2.0	243868	2	AC094180
c	94	26	2.0	245267	2	AC106568
	95	26	2.0	255156	2	AC111130
	96	26	2.0	260876	2	AC127709
	97	26	2.0	270618	2	AC133365
c	98	26	2.0	312459	2	AL928900
	99	25	1.9	62	9	S52152S31
c	100	25	1.9	191	6	AX332084
c	101	25	1.9	191	6	AX333538
c	102	25	1.9	191	6	AX407878
c	103	25	1.9	222	6	AX407940
	104	25	1.9	252	3	AF215118
c	105	25	1.9	330	6	AX255913
	106	25	1.9	365	9	AF251319
	107	25	1.9	390	6	AX645775
	108	25	1.9	430	9	BC058034
	109	25	1.9	438	11	HS542C5S
	110	25	1.9	444	10	BC059153
	111	25	1.9	452	9	AF487338
	112	25	1.9	469	6	AR019396
	113	25	1.9	469	6	AR028949
	114	25	1.9	469	6	AR037083
	115	25	1.9	469	6	AR049639
	116	25	1.9	469	6	AR145297
	117	25	1.9	538	9	BC003669
	118	25	1.9	540	5	AY168623
	119	25	1.9	612	9	BC056906
c	120	25	1.9	657	6	AX525532
	121	25	1.9	671	10	BC027578
	122	25	1.9	704	11	BV023499
	123	25	1.9	716	6	BD216887
	124	25	1.9	729	8	ATAPT
c	125	25	1.9	748	6	BD019727
c	126	25	1.9	748	6	BD099665
	127	25	1.9	775	11	BV076062
	128	25	1.9	780	11	BV038342
	129	25	1.9	836	3	AF457549
	130	25	1.9	876	10	BC008275
	131	25	1.9	893	9	BC009228
	132	25	1.9	938	3	AB072733
	133	25	1.9	953	9	BC011751
	134	25	1.9	1073	10	BC049992
	135	25	1.9	1091	9	BC000566
	136	25	1.9	1149	10	BC022645
	137	25	1.9	1171	9	HSMB00589
	138	25	1.9	1182	10	BC049567

BC056981 Mus muscu
BC056365 Mus muscu
AX059541 Sequence
AB012241 Arabidops
AC010222 Homo sapi
AC007087 Arabidops
AC097626 Homo sapi
AC026773 Homo sapi
AC102011 Mus muscu
AC102011 Mus muscu
AC012477 Genomic S
AC008541 Homo sapi
AC079763 Homo sapi
AC136857 Rattus no
AC012392 Genomic S
AC120345 Mus muscu
AC139042 Mus muscu
BX545915 Danio rer
AL645972 Mouse DNA
AL591146 Mouse DNA
AC131241 Mus muscu
AL161503 Arabidops
AC113953 Mus muscu
AC122268 Mus muscu
AC119244 Mus muscu
AC117813 Mus muscu
AC131321 Mus muscu
AC094180 Rattus no
AC106568 Rattus no
AC111130 Mus muscu
AC127709 Rattus no
AC133365 Rattus no
AL928900 Danio rer
S52228 CD11b=leuko
AX332084 Sequence
AX333538 Sequence
AX407878 Sequence
AX407940 Sequence
AF215118 Conus are
AX255913 Sequence
AF251319 Macaca mu
AX645775 Sequence
BC058034 Homo sapi
AL023956 H.sapiens
BC059153 Rattus no
AF487338 Homo sapi
AR019396 Sequence
AR028949 Sequence
AR037083 Sequence
AR049639 Sequence
AR145297 Sequence
BC003669 Homo sapi
AY168623 Danio rer
BC056906 Homo sapi
AX525532 Sequence
BC027578 Mus muscu
BV023499 S212P6409
BD216887 Novel hum
X58640 A.thaliana
BD019727 Novel gen
BD099665 Novel gen
BV076062 S212P6049
BV038342 S212P6390
AF457549 Anopheles
BC008275 Mus muscu
BC009228 Homo sapi
AB072733 Hemicentr
BC011751 Homo sapi
BC049992 Mus muscu
BC000566 Homo sapi
BC022645 Mus muscu
AL080088 Homo sapi
BC049567 Mus muscu

139	25	1.9	1214	9	AK026312
140	25	1.9	1226	9	BC039664
141	25	1.9	1230	10	BC060975
142	25	1.9	1238	10	BC052530
143	25	1.9	1278	5	AF100931
144	25	1.9	1280	17	AF130066
145	25	1.9	1299	9	BC026126
146	25	1.9	1330	9	BC012495
147	25	1.9	1357	9	BC020805
148	25	1.9	1417	9	BC029650
149	25	1.9	1483	5	AY120892
150	25	1.9	1502	9	BC008376

AK026312 Homo sapi
BC039664 Homo sapi
BC060975 Mus muscu
BC052530 Mus muscu
AF100931 Oncorhync
Af130066 Homo sapi
BC026126 Homo sapi
BC012495 Homo sapi
BC020805 Homo sapi
BC029650 Homo sapi
AY120892 Danio rer
BC008376 Homo sapi

ALIGNMENTS

RESULT 1
BC001222
LOCUS BC001222 3259 bp mRNA linear PRI 04-OCT-2003
DEFINITION Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone MGC:938 IMAGE:3355572), complete cds.
ACCESSION BC001222
VERSION BC001222.1 GI:12654758
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3259)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 3259)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Sco Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 6 Row: j Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8922647.

FEATURES Location/Qualifiers

source 1. .3259
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:938 IMAGE:3355572"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

gene 1. .3259
/gene="FLJ10759"
/db_xref="LocusID:55223"

CDS 34. .1461
/codon_start=1
/product="hypothetical protein FLJ10759"
/protein_id="AAH01222.1"
/db_xref="GI:12654759"
/db_xref="LocusID:55223"
/translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLF
CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
SYFSPGQSHANGKNVQPLRINTVRI"

misc_feature 64. .192
/note="RING; Region: Ring finger"
/db_xref="CDD:smart00184"

misc_feature 445. .798
/note="BBC; Region: B-Box C-terminal domain"
/db_xref="CDD:smart00502"

misc_feature 913. .1074
/note="PRY; Region: associated with SPRY domains"
/db_xref="CDD:smart00589"

misc_feature 1075. .1413
/note="SPRY; Region: Domain in SPla and the RYanodine Receptor."
/db_xref="CDD:smart00449"

ORIGIN

Query Match 100.0%; Score 1327; DB 9; Length 3259;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
Db    1928 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1987

Qy       61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
Db    1988 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2047

Qy      121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
Db    2048 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2107

Qy      181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
Db    2108 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2167

Qy      241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
Db    2168 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2227

Qy      301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
Db    2228 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2287

Qy      361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
Db    2288 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2347

Qy      421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
Db    2348 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2407

Qy      481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
Db    2408 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2467

Qy      541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
Db    2468 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2527

Qy      601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
Db    2528 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2587

Qy      661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
Db    2588 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2647

Qy      721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
Db    2648 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2707

Qy      781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
Db    2708 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2767

Qy      841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
Db    2768 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 2827

Qy      901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
Db    2828 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 2887

Qy      961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
Db    2888 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2947

Qy     1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
Db    2948 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3007

Qy     1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
Db    3008 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3067

Qy     1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
Db    3068 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3127

Qy     1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
Db    3128 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3187

Qy     1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
Db    3188 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3247

Qy     1321 AAAAAAA 1327
```

```
                    |||||||
Db       3248 AAAAAAA 3254


RESULT 2
BC007999
LOCUS       BC007999               3259 bp    mRNA    linear   PRI 03-OCT-2003
DEFINITION  Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone
            MGC:15757 IMAGE:3357436), complete cds.
ACCESSION   BC007999
VERSION     BC007999.1  GI:14124949
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3259)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3259)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 23 Row: a Column: 20.
FEATURES             Location/Qualifiers
     source          1. .3259
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:15757 IMAGE:3357436"
                     /tissue_type="Eye, retinoblastoma"
                     /clone_lib="NIH_MGC_16"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .3259
                     /gene="FLJ10759"
                     /db_xref="LocusID:55223"
     CDS             34. .1461
                     /codon_start=1
                     /product="hypothetical protein FLJ10759"
                     /protein_id="AAH07999.1"
                     /db_xref="GI:14124950"
                     /db_xref="LocusID:55223"
                     /translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
                     QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLF
                     CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
                     KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
                     SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
                     PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
                     DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
                     DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
                     SYFSPGQSHANGKNVQPLRINTVRI"
     misc_feature    64. .192
                     /note="RING; Region: Ring finger"
                     /db_xref="CDD:smart00184"
     misc_feature    445. .798
                     /note="BBC; Region: B-Box C-terminal domain"
                     /db_xref="CDD:smart00502"
     misc_feature    913. .1074
                     /note="PRY; Region: associated with SPRY domains"
                     /db_xref="CDD:smart00589"
     misc_feature    1075. .1413
                     /note="SPRY; Region: Domain in SPla and the RYanodine
                     Receptor."
                     /db_xref="CDD:smart00449"
ORIGIN

  Query Match           100.0%;  Score 1327;  DB 9;  Length 3259;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1928 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1987

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTCAGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1988 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTCAGAT 2047

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2048 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2107

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2108 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2167

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2168 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2227

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2228 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2287

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2288 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2347

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2348 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2407
```

```
Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2408 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2467

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2468 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2527

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2528 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2587

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2588 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2647

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2648 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2707

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2708 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2767

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2768 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 2827

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2828 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 2887

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2888 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2947

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2948 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3007

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3008 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3067

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3068 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3127

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3128 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3187

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3247

Qy       1321 AAAAAAA 1327
              |||||||
Db       3248 AAAAAAA 3254
```

```
RESULT 3
BC011689
LOCUS       BC011689                3259 bp    mRNA    linear   PRI 19-DEC-2003
DEFINITION  Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone
            MGC:19672 IMAGE:3353034), complete cds.
ACCESSION   BC011689
VERSION     BC011689.2  GI:40225630
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3259)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE   2  (bases 1 to 3259)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Dec 19, 2003 this sequence version replaced gi:15079757.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 27 Row: b Column: 5
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES                Location/Qualifiers
     source             1. .3259
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:19672 IMAGE:3353034"
                        /tissue_type="Eye, retinoblastoma"
                        /clone_lib="NIH_MGC_16"
                        /lab_host="DH10B-R"
                        /note="Vector: pOTB7"
     gene               1. .3259
                        /gene="FLJ10759"
                        /db_xref="LocusID:55223"
     CDS                34. .1461
                        /gene="FLJ10759"
                        /codon_start=1
```

```
                     /product="hypothetical protein FLJ10759"
                     /protein_id="AAH11689.1"
                     /db_xref="GI:15079758"
                     /db_xref="LocusID:55223"
                     /translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
                     QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLF
                     CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
                     KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
                     SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
                     PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
                     DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
                     DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
                     SYFSPGQSHANGKNVQPLRINTVRI"
     misc_feature    64. .204
                     /gene="FLJ10759"
                     /note="RING; Region: RING-finger (Really Interesting New
                     Gene) domain, a specialized type of Zn-finger of 40 to 60
                     residues that binds two atoms of zinc"
                     /db_xref="CDD:cd00162"
     misc_feature    445. .798
                     /gene="FLJ10759"
                     /note="BBC; Region: B-Box C-terminal domain"
                     /db_xref="CDD:smart00502"
     misc_feature    913. .1074
                     /gene="FLJ10759"
                     /note="PRY; Region: associated with SPRY domains"
                     /db_xref="CDD:smart00589"
     misc_feature    1075. .1413
                     /gene="FLJ10759"
                     /note="SPRY; Region: Domain in SPla and the RYanodine
                     Receptor."
                     /db_xref="CDD:smart00449"
ORIGIN

  Query Match           100.0%;  Score 1327;  DB 9;  Length 3259;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1928 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1987

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1988 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2047

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2048 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2107

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2108 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2167

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2168 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2227

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2228 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2287

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2288 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2347

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2348 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2407

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2408 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2467

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2468 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2527

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2528 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2587

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2588 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2647

Qy        721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2648 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2707

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2708 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2767

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2768 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 2827

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2828 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 2887

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2888 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2947

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2948 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3007

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3008 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3067

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3068 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3127

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3128 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3187

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3247

Qy       1321 AAAAAAA 1327
              |||||||
Db       3248 AAAAAAA 3254


RESULT 4
BC012152
LOCUS       BC012152                3259 bp    mRNA    linear   PRI 04-OCT-2003
DEFINITION  Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone
            MGC:20370 IMAGE:4558639), complete cds.
ACCESSION   BC012152
VERSION     BC012152.1  GI:15082475
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3259)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
```

```
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3259)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 29 Row: d Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 8922647.
FEATURES            Location/Qualifiers
     source         1. .3259
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="MGC:20370 IMAGE:4558639"
                    /tissue_type="Eye, retinoblastoma"
                    /clone_lib="NIH_MGC_16"
                    /lab_host="DH10B-R"
                    /note="Vector: pOTB7"
     gene           1. .3259
                    /gene="FLJ10759"
                    /db_xref="LocusID:55223"
     CDS            34. .1461
                    /codon_start=1
                    /product="hypothetical protein FLJ10759"
                    /protein_id="AAH12152.1"
                    /db_xref="GI:15082476"
                    /db_xref="LocusID:55223"
                    /translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
                    QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLF
                    CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
                    KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
                    SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
                    PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
                    DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
                    DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
                    SYFSPGQSHANGKNVQPLRINTVRI"
     misc_feature   64. .192
                    /note="RING; Region: Ring finger"
                    /db_xref="CDD:smart00184"
     misc_feature   445. .798
                    /note="BBC; Region: B-Box C-terminal domain"
                    /db_xref="CDD:smart00502"
     misc_feature   913. .1074
                    /note="PRY; Region: associated with SPRY domains"
                    /db_xref="CDD:smart00589"
     misc_feature   1075. .1413
                    /note="SPRY; Region: Domain in SPla and the RYanodine
                    Receptor."
                    /db_xref="CDD:smart00449"
ORIGIN

  Query Match             100.0%;  Score 1327;  DB 9;  Length 3259;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1928 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1987

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1988 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2047

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2048 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2107

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2108 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2167

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2168 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2227

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2228 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2287

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2288 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2347

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2348 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2407

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2408 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2467

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2468 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2527

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2528 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2587

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2588 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2647
```

```
Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2648 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2707

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2708 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2767

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2768 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 2827

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2828 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 2887

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2888 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2947

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2948 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3007

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3008 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3067

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3068 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3127

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3128 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3187

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3247

Qy       1321 AAAAAAA 1327
              |||||||
Db       3248 AAAAAAA 3254
```

```
RESULT 5
AX491279
LOCUS       AX491279               3826 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 3 from Patent WO0212285.
ACCESSION   AX491279
VERSION     AX491279.1  GI:22323980
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Killary,A., Chandler,D. and Lott,S.
  TITLE     The tumor suppressor car-1
  JOURNAL   Patent: WO 0212285-A 3 14-FEB-2002;
            BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)
FEATURES             Location/Qualifiers
     source          1. .3826
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 1327;  DB 6;  Length 3826;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2559

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2560 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2619

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2679

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2680 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2739

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2740 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2799

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2800 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2859

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2860 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2919

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2920 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2979

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2980 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 3039

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3040 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 3099

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3100 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 3159

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3160 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 3219

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3220 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 3279

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3280 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 3339

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3340 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3399

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3400 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3459

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3460 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3519

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3520 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3579

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3580 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3639

Qy       1141 AATGGCTGCCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3640 AATGGCTGCCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3699

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3700 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3759

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3760 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3819

Qy       1321 AAAAAAA 1327
              |||||||
Db       3820 AAAAAAA 3826


RESULT 6
AX491283
LOCUS       AX491283               23433 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 7 from Patent WO0212285.
ACCESSION   AX491283
VERSION     AX491283.1  GI:22323984
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Killary,A., Chandler,D. and Lott,S.
  TITLE     The tumor suppressor car-1
  JOURNAL   Patent: WO 0212285-A 7 14-FEB-2002;
            BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)
FEATURES             Location/Qualifiers
     source          1. .23433
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             98.6%;  Score 1308;  DB 6;  Length 23433;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1308;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13550 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 13609

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13610 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 13669

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13670 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 13729

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13730 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 13789

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13790 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 13849

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13850 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 13909

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13910 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 13969

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13970 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 14029

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14030 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 14089

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14090 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 14149

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14150 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 14209

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14210 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 14269

Qy        721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14270 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 14329

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14330 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 14389

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14390 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 14449

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14450 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 14509

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14510 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 14569

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14570 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 14629

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14630 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 14689

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14690 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 14749

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14750 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 14809

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      14810 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 14857


RESULT 7
AC022262/c
LOCUS       AC022262               164950 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-150F21, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC022262
VERSION     AC022262.4  GI:8569825
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164950)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164950)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 16, 2000 this sequence version replaced gi:7321934.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0150F21
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 75%
            Sequencing vector: plasmid; 25%
            Chemistry: Dye-primer ET; 75% of reads
            Chemistry: Dye-terminator Big Dye; 25% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 163188 bases at least Q40
            Consensus quality: 163696 bases at least Q30
            Consensus quality: 163935 bases at least Q20
            Insert size: 172000; agarose-fp
            Insert size: 164650; sum-of-contigs
            Quality coverage: 7.83 in Q20 bases; agarose-fp
            Quality coverage: 8.20 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     8662: contig of 8662 bp in length
            *     8663     8762: gap of unknown length
            *     8763    27020: contig of 18258 bp in length
            *    27021    27120: gap of unknown length
            *    27121    78035: contig of 50915 bp in length
            *    78036    78135: gap of unknown length
            *    78136   164950: contig of 86815 bp in length.
FEATURES             Location/Qualifiers
     source          1. .164950
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-150F21"
     misc_feature    1. .8662
                     /note="assembly_name:Contig3
                     clone_end:SP6
                     vector_side:left"
     misc_feature    8763. .27020
                     /note="assembly_name:Contig4"
     misc_feature    27121. .78035
                     /note="assembly_name:Contig5
                     clone_end:T7
                     vector_side:right"
     misc_feature    78136. .164950
                     /note="assembly_name:Contig6"
ORIGIN

  Query Match             98.6%;  Score 1308;  DB 2;  Length 164950;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1308;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18641 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 18582

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTCAGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18581 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTCAGAT 18522

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18521 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 18462

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18461 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 18402

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18401 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 18342

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18341 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 18282

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18281 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 18222

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18221 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 18162

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18161 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 18102

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18101 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 18042

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18041 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 17982

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17981 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 17922

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17921 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 17862

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17861 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 17802

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17801 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 17742

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17741 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 17682

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17681 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 17622

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17621 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 17562

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
```

```
Db      17561 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 17502

Qy       1141 AATGGCTGCCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
Db      17501 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 17442

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
Db      17441 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 17382

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
Db      17381 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 17334


RESULT 8
AX775831
LOCUS       AX775831               3243 bp    mRNA    linear   PAT 14-JUL-2003
DEFINITION  Sequence 101 from Patent WO03048202.
ACCESSION   AX775831
VERSION     AX775831.1  GI:32693549
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Matsuda,A. and Muramatsu,S.
  TITLE     NF-kB activating gene
  JOURNAL   Patent: WO 03048202-A 101 12-JUN-2003;
            Asahi Kasei Kabushiki Kaisha (JP)
FEATURES             Location/Qualifiers
     source          1. .3243
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             43. .1470
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE11615.1"
                     /db_xref="GI:32693550"
                     /translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
                     QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLF
                     CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
                     KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
                     SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
                     PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
                     DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
                     DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
                     SYFSPGQSHANGKNVQPLRINTVRI"
ORIGIN

  Query Match           98.5%;  Score 1307;  DB 6;  Length 3243;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1307;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
Db       1937 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1996

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
Db       1997 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2056

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
Db       2057 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2116

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
Db       2117 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2176

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
Db       2177 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2236

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
Db       2237 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2296

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
Db       2297 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2356

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
Db       2357 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2416

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
Db       2417 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2476

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
Db       2477 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2536

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
Db       2537 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2596

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
Db       2597 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2656

Qy        721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
Db       2657 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2716

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
Db       2717 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2776

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
Db       2777 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 2836

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
Db       2837 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 2896

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
Db       2897 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2956

Qy       1021 AATGTCTCCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
Db       2957 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3016

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
Db       3017 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3076

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
Db       3077 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3136

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
Db       3137 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3196

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
Db       3197 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 3243
```

```
RESULT 9
AX491284
LOCUS       AX491284               30676 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 8 from Patent WO0212285.
ACCESSION   AX491284
VERSION     AX491284.1  GI:22323985
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Killary,A., Chandler,D. and Lott,S.
  TITLE     The tumor suppressor car-1
  JOURNAL   Patent: WO 0212285-A 8 14-FEB-2002;
            BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)
FEATURES             Location/Qualifiers
     source          1. .30676
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             94.7%;  Score 1257;  DB 6;  Length 30676;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1307;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24440 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 24499

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24500 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 24559

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24560 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 24619

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24620 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 24679

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24680 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 24739

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24740 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 24799

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24800 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 24859

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24860 GGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 24919

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24920 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 24979

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24980 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 25039

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25040 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 25099

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      25100 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 25159

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25160 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 25219

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25220 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 25279

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25280 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 25339

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25340 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 25399

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25400 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 25459

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25460 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 25519

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25520 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 25579

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25580 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 25639

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25640 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 25699

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      25700 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 25747


RESULT 10
AL662907/c
LOCUS       AL662907               64693 bp    DNA     linear   PRI 19-JUN-2002
DEFINITION  Human DNA sequence from clone RP11-131M11 on chromosome 1, complete
            sequence.
ACCESSION   AL662907 AC026053
VERSION     AL662907.11  GI:21531085
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 64693)
  AUTHORS   Van Hellmond,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 21, 2002 this sequence version replaced gi:20145349.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
```

```
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-131M11 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            Draft Sequence Produced by Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            http://genome.wustl.edu/gsc/index.shtml.
FEATURES             Location/Qualifiers
     source          1. .64693
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-131M11"
                     /clone_lib="RPCI-11.1"
ORIGIN

  Query Match             94.7%;  Score 1257;  DB 9;  Length 64693;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1307;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7232 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 7173

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7172 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 7113

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7112 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 7053

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7052 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 6993

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6992 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 6933

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6932 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 6873

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6872 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 6813

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6812 GGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 6753

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6752 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 6693

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6692 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 6633

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6632 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 6573

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6572 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 6513

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6512 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 6453

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6452 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 6393

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6392 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 6333

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6332 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 6273

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6272 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 6213

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6212 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 6153

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6152 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 6093

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6092 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 6033

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6032 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 5973

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       5972 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 5925


RESULT 11
AX877131
LOCUS       AX877131               3436 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 12036 from Patent EP1074617.
ACCESSION   AX877131
VERSION     AX877131.1  GI:40031867
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 12036 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .3436
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             235. .1662
                     /note="unnamed protein product"
```

```
                     /codon_start=1
                     /protein_id="CAE89585.1"
                     /db_xref="GI:40031868"
                     /translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
                     QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLL
                     CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
                     KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
                     SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
                     PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
                     DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
                     DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
                     SYFSPGQSHANGKNVQPLRINTVLI"
ORIGIN

  Query Match             94.6%;  Score 1256;  DB 6;  Length 3436;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1306;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2189

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2249

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2250 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2309

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2310 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2369

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2370 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2429

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2430 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2489

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2490 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2549

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2550 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2609

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2610 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2669

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2670 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2729

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2730 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2789

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2849

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2909

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2969

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 CATCTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3029

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3089

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3149

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3209

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3269

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3329

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3389

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 3436


RESULT 12
BD156501
LOCUS       BD156501                3436 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD156501
VERSION     BD156501.1  GI:27862259
KEYWORDS    JP 2002191363-A/11344.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3436)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11344 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11344
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                       Location/Qualifiers
            FT   CDS             (235). .(1659).
FEATURES                Location/Qualifiers
     source             1. .3436
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN
```

```
Query Match             94.6%;  Score 1256;  DB 6;  Length 3436;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 1306;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2189

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2249

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2250 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2309

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2310 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2369

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2370 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2429

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2430 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2489

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2490 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2549

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2550 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2609

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2610 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2669

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2670 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2729

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2730 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2789

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2849

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2909

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2969

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 CATCTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3029

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3089

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3149

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3209

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3269

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3329

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3389

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 3436
```

RESULT 13
AK001621
LOCUS AK001621 3436 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ10759 fis, clone NT2RP3004617, weakly similar to ZINC-BINDING PROTEIN A33.
ACCESSION AK001621
VERSION AK001621.1 GI:7022986
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Kanehori,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3436)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES Location/Qualifiers
source 1. .3436
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3004617"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP3"
/note="cloning vector: pME18SFL3~mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
CDS 235. .1662
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91792.1"
/db_xref="GI:7022987"
/translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLL CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL

KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQPSRGFYCIVMH
DGNQYSACTEPWTRLNVRDKLDKVGVFLDYDQGLLIFYNADDMSWLYTFREKFPGKLC
SYFSPGQSHANGKNVQPLRINTVLI"

ORIGIN

Query Match 94.6%; Score 1256; DB 9; Length 3436;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1306; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2189

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2249

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2250 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2309

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2310 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2369

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2370 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2429

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2430 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2489

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2490 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2549

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2550 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2609

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2610 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2669

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2670 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2729

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2730 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2789

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2849

Qy        721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2909

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2969

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 CATCTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3029

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3089

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3149

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3209

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3269

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3329

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3389

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 3436
```

RESULT 14
AF161326
LOCUS AF161326 136095 bp DNA linear HTG 26-JUL-2002
DEFINITION Homo sapiens chromosome 8 clone CTA-392H5 map 8q11, WORKING DRAFT SEQUENCE, 12 unordered pieces.
ACCESSION AF161326
VERSION AF161326.3 GI:16356844
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 136095)
AUTHORS Wen,G., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N., Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A., Siddiqui,R., Taudien,S., Schilhabel,M.B., Schlegelberger,B., Siebert,R., Rosenthal,A. and Platzer,M.
TITLE Chromosome 8 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 136095)
AUTHORS Polley,A., Nordsiek,G., Schlegelberger,B., Drescher,B., Weber,J., Schattevoy,R. and Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-1999) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT On Oct 24, 2001 this sequence version replaced gi:14277237.
---------- Genome Center
Center: Insitute of Molecular Biotechnoloy
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
---------- Project Information
Center project name: H176
Center clone name: CTA-392H5
---------- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 122353 bases at least Q40
Consensus quality: 127794 bases at least Q30
Consensus quality: 131293 bases at least Q20
Quality coverage: 5.82 x in Q20 bases; sum-of-contigs
----------.
Sequence Quality Assessment:
This entry has been annotated with sequence quality

```
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality 10.
            Quality levels above 40 are expected to have less than
            1 error in 10,000 bp.
            Base-by-base quality values are not generally visible from the
            GenBank flat file format but are available as part
            of this entry's ASN.1 file.
            ----------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1802: contig of 1802 bp in length
            *     1803     1902: gap of unknown length
            *     1903     5293: contig of 3391 bp in length
            *     5294     5393: gap of unknown length
            *     5394     9974: contig of 4581 bp in length
            *     9975    10074: gap of unknown length
            *    10075    15065: contig of 4991 bp in length
            *    15066    15165: gap of unknown length
            *    15166    20118: contig of 4953 bp in length
            *    20119    20218: gap of unknown length
            *    20219    26289: contig of 6071 bp in length
            *    26290    26389: gap of unknown length
            *    26390    35964: contig of 9575 bp in length
            *    35965    36064: gap of unknown length
            *    36065    46363: contig of 10299 bp in length
            *    46364    46463: gap of unknown length
            *    46464    58765: contig of 12302 bp in length
            *    58766    58865: gap of unknown length
            *    58866    77219: contig of 18354 bp in length
            *    77220    77319: gap of unknown length
            *    77320   101085: contig of 23766 bp in length
            *   101086   101185: gap of unknown length
            *   101186   136095: contig of 34910 bp in length.
FEATURES             Location/Qualifiers
     source          1. .136095
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q11"
                     /clone="CTA-392H5"
ORIGIN

  Query Match             79.4%;  Score 1053;  DB 2;  Length 136095;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1303;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67212 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 67271

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67272 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 67331

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67332 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 67391

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      67392 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAAAG 67451

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67452 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 67511

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67512 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 67571

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67572 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 67631

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67632 AAACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 67691

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67692 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 67751

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67752 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 67811

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67812 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 67871

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67872 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 67931

Qy        721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67932 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 67991

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67992 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 68051

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68052 CATTTCCCTAAAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 68111

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68112 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 68171

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68172 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 68231

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      68232 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAAATTCTGCACTTGGGG 68291

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      68292 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAAAGCTAAACACAGGGGGTTCTTAA 68351

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68352 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 68411

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68412 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 68471

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      68472 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 68519

RESULT 15
AX491281
LOCUS       AX491281               30625 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 5 from Patent WO0212285.
```

```
ACCESSION   AX491281
VERSION     AX491281.1  GI:22323982
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Killary,A., Chandler,D. and Lott,S.
  TITLE     The tumor suppressor car-1
  JOURNAL   Patent: WO 0212285-A 5 14-FEB-2002;
            BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)
FEATURES             Location/Qualifiers
     source          1. .30625
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           54.7%;  Score 726;  DB 6;  Length 30625;
  Best Local Similarity 99.4%;  Pred. No. 0;
  Matches 1146;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        157 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23131 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 23190

Qy        217 GGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 276
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      23191 GGTATACCCAAAGCTGATGCAAAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 23250

Qy        277 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23251 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 23310

Qy        337 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23311 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 23370

Qy        397 AAGCACT-GGGAGGGGCTGTTGGCTAGACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 455
              ||||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      23371 AAGCACTGGGGAGGGGCTGTTGGCTAAACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 23430

Qy        456 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23431 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 23490

Qy        516 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23491 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 23550

Qy        576 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23551 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 23610

Qy        636 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23611 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 23670

Qy        696 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCC 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23671 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCC 23730

Qy        756 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23731 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 23790

Qy        816 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23791 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 23850

Qy        876 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGA 935
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      23851 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAAAACTCGGGAGAGGGTGGA 23910

Qy        936 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23911 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 23970

Qy        996 GGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 1055
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      23971 GGAGTGGGCTGCAAAACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 24030

Qy       1056 TGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 1115
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24031 TGGGGCCTAAATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 24090

Qy       1116 AGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 1175
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24091 AAAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 24150

Qy       1176 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24151 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 24210

Qy       1236 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24211 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 24270

Qy       1296 GTTACAGATGTCA 1308
              |||||||||||||
Db      24271 GTTACAGATGTCA 24283


RESULT 16
AX775827
LOCUS       AX775827               2207 bp    mRNA    linear   PAT 14-JUL-2003
DEFINITION  Sequence 97 from Patent WO03048202.
ACCESSION   AX775827
VERSION     AX775827.1  GI:32693545
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Matsuda,A. and Muramatsu,S.
  TITLE     NF-kB activating gene
  JOURNAL   Patent: WO 03048202-A 97 12-JUN-2003;
            Asahi Kasei Kabushiki Kaisha (JP)
FEATURES             Location/Qualifiers
     source          1. .2207
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             138. .446
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE11613.1"
                     /db_xref="GI:32693546"
                     /translation="MYGSAWLPPPPGFPIPSFSPSPSPHRPHPGISTPLRAPPRSGIPFS
                     QLLSLRTAPPPRTSPSGSPSLSRSQPPPSAHRLELPTTSRLPSRAGVPRSLSRAPLLR
                     "
ORIGIN

  Query Match           34.1%;  Score 452;  DB 6;  Length 2207;
  Best Local Similarity 99.8%;  Pred. No. 1.6e-263;
  Matches  502;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        801 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1705 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 1764

Qy        861 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC 920
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1765 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTCTTCTGCCACAGCATCCAATCGTGAAGAAC 1824
```

```
Qy        921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1825 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 1884

Qy        981 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1885 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1944

Qy       1041 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1945 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 2004

Qy       1101 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2005 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 2064

Qy       1161 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2065 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 2124

Qy       1221 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2125 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 2184

Qy       1281 TTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||
Db       2185 TTGTATTAAATTAAAGTTACAGA 2207
```

RESULT 17
AX776017
LOCUS AX776017 2207 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 287 from Patent WO03048202.
ACCESSION AX776017
VERSION AX776017.1 GI:32693735
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Matsuda,A. and Muramatsu,S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 287 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES Location/Qualifiers
source 1. .2207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
CDS 571. .1755
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11708.1"
/db_xref="GI:32693736"
/translation="MACSLKDELLCSICLSIYQDPVSLGCEHYFCRRCITEHWVRQEA
QGARDCPECRRTFAEPALAPSLKLANIVERYSSFPLDAILNARRAARPCQAHDKVKLF
CLTDRALLCFFCDEPALHEQHQVTGIDDAFDELQRELKDQLQALQDSEREHTEALQLL
KRQLAETKSSTKSLRTTIGEAFERLHRLLRERQKAMLEELEADTARTLTDIEQKVQRY
SQQLRKVQEGAQILQERLAETDRHTFLAGVASLSERLKGKIHETNLTYEDFPTSKYTG
PLQYTIWKSLFQDIHPVPAALTLDPGTAHQRLILSDDCTIVAYGNLHPQPLQDSPKRF
DVEVSVLGSEAFSSGVHYWEVVVAEKTQWVIGLAHEAASRKGSIQIQNLGCKGVWEMS
FP"
ORIGIN

```
  Query Match            34.1%;  Score 452;  DB 6;  Length 2207;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-263;
  Matches  502;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        801 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1705 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 1764

Qy        861 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1765 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTCTTCTGCCACAGCATCCAATCGTGAAGAAC 1824

Qy        921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1825 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 1884

Qy        981 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1885 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1944

Qy       1041 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1945 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 2004

Qy       1101 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2005 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 2064

Qy       1161 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2065 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 2124

Qy       1221 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2125 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 2184

Qy       1281 TTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||
Db       2185 TTGTATTAAATTAAAGTTACAGA 2207
```

RESULT 18
AX775829
LOCUS AX775829 2246 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 99 from Patent WO03048202.
ACCESSION AX775829
VERSION AX775829.1 GI:32693547
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Matsuda,A. and Muramatsu,S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 99 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES Location/Qualifiers
source 1. .2246
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
CDS 138. .485
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11614.1"
/db_xref="GI:32693548"
/translation="MYGSAWLPPPPGFPIPSFSPSPHRPHPGISTPLRAPPRSGIPFS
QLLSLRTAPPPRTSPSGSPSLSRSQPPPSAHRLELPTTSRLPSRAGVPRSLSRAPLPR
VTLLPDSTPSLLR"
ORIGIN

```
  Query Match            34.1%;  Score 452;  DB 6;  Length 2246;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-263;
  Matches  502;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        801 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 1803
```

```
Qy        861 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC 920
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1804 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTCTTCTGCCACAGCATCCAATCGTGAAGAAC 1863

Qy        921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 1923

Qy        981 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1983

Qy       1041 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 2043

Qy       1101 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 2103

Qy       1161 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2104 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 2163

Qy       1221 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2164 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 2223

Qy       1281 TTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||
Db       2224 TTGTATTAAATTAAAGTTACAGA 2246
```

RESULT 19
AX873748/c
LOCUS AX873748 557 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 8653 from Patent EP1074617.
ACCESSION AX873748
VERSION AX873748.1 GI:40028540
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 8653 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES Location/Qualifiers
source 1. .557
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 31.9%; Score 423; DB 6; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.2e-246;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        885 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 364

Qy        945 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 304

Qy       1005 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 244

Qy       1065 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 184

Qy       1125 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 124

Qy       1185 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 64

Qy       1245 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 4

Qy       1305 GTC 1307
              |||
Db           3 GTC 1
```

RESULT 20
BD153810/c
LOCUS BD153810 557 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD153810
VERSION BD153810.1 GI:27859568
KEYWORDS JP 2002191363-A/8653.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 557)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 8653 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/8653
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1. .557
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .557
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 31.9%; Score 423; DB 6; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.2e-246;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        885 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 364

Qy        945 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 304
```

```
Qy       1005 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 1064
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 244

Qy       1065 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 1124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 184

Qy       1125 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 1184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 124

Qy       1185 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 1244
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 64

Qy       1245 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 1304
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 4

Qy       1305 GTC 1307
             |||
Db           3 GTC 1
```

RESULT 21
AK122896

```
LOCUS       AK122896                3148 bp    mRNA    linear   PRI 09-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ16558 fis, clone SYNOV4001153, weakly similar
            to Zinc-binding protein A33.
ACCESSION   AK122896
VERSION     AK122896.1  GI:34528340
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3148)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .3148
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SYNOV4001153"
                     /tissue_type="synovial membrane tissue from rheumatoid
                     arthritis"
                     /clone_lib="SYNOV4"
                     /note="cloning vector: pME18SFL3"
ORIGIN
```

```
  Query Match             25.3%;  Score 336;  DB 9;  Length 3148;
  Best Local Similarity   100.0%;  Pred. No. 1e-192;
  Matches  336;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        965 GCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATG 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATG 2872

Qy       1025 TCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTC 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTC 2932

Qy       1085 TGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATG 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 TGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATG 2992

Qy       1145 GCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCC 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 GCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCC 3052

Qy       1205 TTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTT 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 TTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTT 3112

Qy       1265 ATGAGTTATGATCAATTTGTATTAAATTAAAGTTAC 1300
              ||||||||||||||||||||||||||||||||||||
Db       3113 ATGAGTTATGATCAATTTGTATTAAATTAAAGTTAC 3148
```

RESULT 22
G15084/c

```
LOCUS       G15084                   250 bp    DNA     linear   STS 04-JAN-1996
DEFINITION  human STS SHGC-15366, sequence tagged site.
ACCESSION   G15084
VERSION     G15084.1  GI:1131847
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 250)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: AGTTGAAGGGTTGGGGTAGG
            Primer B: CACTTGGGGTCTCTGACACA
            STS size: 140
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:       94 degrees C for 15 seconds
                    Annealing:          62 degrees C for 23 seconds
                    Polymerization:     72 degrees C for 30 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9600
            Protocol:
                    Template:       25 ng
                    Primer:         each 1 uM
                    dNTPs:          each 200 uM
                    Taq Polymerase: 0.05 units/ul
                    Total Vol:      10 ul

            Buffer:
                    MgCl2:          2.5 mM
                    KCl:            50  mM
```

```
                    Tris-HCl:       20  mM
                    pH:             8.3

          Prepared with primer pairs derived from T85347--Merck/UniEST.
FEATURES             Location/Qualifiers
     source          1. .250
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="1"
     STS             109. .248
     primer_bind     109. .128
     primer_bind     complement(229. .248)
ORIGIN

  Query Match              4.7%;  Score 62;  DB 11;  Length 250;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-25;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1213 CCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTA 1272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTA 48

Qy       1273 TG 1274
             ||
Db         47 TG 46


RESULT 23
AX388847
LOCUS       AX388847                 614 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 3775 from Patent WO0214500.
ACCESSION   AX388847
VERSION     AX388847.1  GI:19581977
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,
            Randazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and
            Labat,I.
  TITLE     Human genes and gene expression products
  JOURNAL   Patent: WO 0214500-A 3775 21-FEB-2002;
            CHIRON CORPORATION (US) ; Hyseq Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match              2.2%;  Score 29;  DB 6;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-05;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||||
Db        585 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 613


RESULT 24
AX771586/c
LOCUS       AX771586                 1652 bp    DNA     linear   PAT 02-JUL-2003
DEFINITION  Sequence 299 from Patent WO03004646.
ACCESSION   AX771586
VERSION     AX771586.1  GI:32438384
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   O'Mahony,D.J., Byrne,D., Brayden,D., Lambkin,I. and Higgins,L.
  TITLE     Genetic analysis of Peyers's patches and M cells and methods and
            compositions targeting Peyer's patches and M cell receptors
  JOURNAL   Patent: WO 03004646-A 299 16-JAN-2003;
            ELAN CORPORATION, Plc (IE)
FEATURES             Location/Qualifiers
     source          1. .1652
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match              2.2%;  Score 29;  DB 6;  Length 1652;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-05;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||||
Db       1474 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1446


RESULT 25
AC019255/c
LOCUS       AC019255                 162959 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 12 clone RP11-133N3 map 12, WORKING DRAFT
            SEQUENCE, 17 unordered pieces.
ACCESSION   AC019255
VERSION     AC019255.3  GI:10045315
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 162959)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 12, clone RP11-133N3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 162959)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 162959)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
            Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
            Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,
            Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
```

```
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
            Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 9, 2000 this sequence version replaced gi:6759192.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2578
                Center clone name: 133_N_3
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 154857 bases at least Q40
                Consensus quality: 158391 bases at least Q30
                Consensus quality: 159656 bases at least Q20
                Insert size: 172000; agarose-fp
                Insert size: 161359; sum-of-contigs
                Quality coverage: 5.1 in Q20 bases; agarose-fp
                Quality coverage: 5.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    30551: contig of 30551 bp in length
            *    30552    30651: gap of 100 bp
            *    30652    33510: contig of 2859 bp in length
            *    33511    33610: gap of 100 bp
            *    33611    36481: contig of 2871 bp in length
            *    36482    36581: gap of 100 bp
            *    36582    39763: contig of 3182 bp in length
            *    39764    39863: gap of 100 bp
            *    39864    45400: contig of 5537 bp in length
            *    45401    45500: gap of 100 bp
            *    45501    49120: contig of 3620 bp in length
            *    49121    49220: gap of 100 bp
            *    49221    53633: contig of 4413 bp in length
            *    53634    53733: gap of 100 bp
            *    53734    58228: contig of 4495 bp in length
            *    58229    58328: gap of 100 bp
            *    58329    63731: contig of 5403 bp in length
            *    63732    63831: gap of 100 bp
            *    63832    69278: contig of 5447 bp in length
            *    69279    69378: gap of 100 bp
            *    69379    76967: contig of 7589 bp in length
            *    76968    77067: gap of 100 bp
            *    77068    86466: contig of 9399 bp in length
            *    86467    86566: gap of 100 bp
            *    86567    97145: contig of 10579 bp in length
            *    97146    97245: gap of 100 bp
            *    97246   110405: contig of 13160 bp in length
            *   110406   110505: gap of 100 bp
            *   110506   135582: contig of 25077 bp in length
            *   135583   135682: gap of 100 bp
            *   135683   161840: contig of 26158 bp in length
            *   161841   161940: gap of 100 bp
            *   161941   162959: contig of 1019 bp in length
FEATURES             Location/Qualifiers
     source          1. .162959
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12"
                     /clone="RP11-133N3"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .30551
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    30652. .33510
                     /note="assembly_fragment"
     misc_feature    33611. .36481
                     /note="assembly_fragment"
     misc_feature    36582. .39763
                     /note="assembly_fragment"
     misc_feature    39864. .45400
                     /note="assembly_fragment"
     misc_feature    45501. .49120
                     /note="assembly_fragment"
     misc_feature    49221. .53633
                     /note="assembly_fragment"
     misc_feature    53734. .58228
                     /note="assembly_fragment"
     misc_feature    58329. .63731
                     /note="assembly_fragment"
     misc_feature    63832. .69278
                     /note="assembly_fragment"
     misc_feature    69379. .76967
                     /note="assembly_fragment"
     misc_feature    77068. .86466
                     /note="assembly_fragment"
     misc_feature    86567. .97145
                     /note="assembly_fragment"
     misc_feature    97246. .110405
                     /note="assembly_fragment"
     misc_feature    110506. .135582
                     /note="assembly_fragment"
     misc_feature    135683. .161840
                     /note="assembly_fragment"
     misc_feature    161941. .162959
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             2.2%;  Score 29;  DB 2;  Length 162959;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-05;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||||
Db     101907 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 101879


RESULT 26
AC022080
LOCUS       AC022080               178217 bp    DNA     linear   PRI 15-MAR-2003
DEFINITION  Homo sapiens 12 BAC RP11-820K3 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC022080
VERSION     AC022080.37  GI:22002110
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 178217)
```

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 178217)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 178217)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 178217)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 178217)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 6 (bases 1 to 178217)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 7 (bases 1 to 178217)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jul 29, 2002 this sequence version replaced gi:21914534.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES Location/Qualifiers
source 1. .178217
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-820K3"
misc_feature complement(1. .2015)
/note="overlaps bases 1. .2004 of clone AC068792"
/function="clone overlap"
repeat_region complement(119. .424)
/rpt_family="AluY"
repeat_region 564. .665
/rpt_family="U13"
repeat_region complement(760. .1052)
/rpt_family="AluJb"
misc_feature 1583
/note="polymorphism G/T AC022080/AC068792"
/function="polymorphic site"
repeat_region complement(1713. .2020)
/rpt_family="AluSg"
misc_feature 1713
/note="polymorphism T/A AC022080/AC068792"
/function="polymorphic site"
STS 2073. .2196
/standard_name="RH92710"
repeat_region 2261. .2296
/rpt_family="AT_rich"

```
     repeat_region   2943. .3072
                     /rpt_family="FLAM_C"
     repeat_region   3743. .3775
                     /rpt_family="AT_rich"
     repeat_region   4258. .4559
                     /rpt_family="AluSx"
     repeat_region   complement(4640. .4683)
                     /rpt_family="MIR"
     repeat_region   complement(4840. .5140)
                     /rpt_family="AluSg1"
     repeat_region   complement(5211. .5517)
                     /rpt_family="AluY"
     repeat_region   6138. .6443
                     /rpt_family="AluSg"
     repeat_region   6446. .6747
                     /rpt_family="AluSx"
     repeat_region   7202. .7489
                     /rpt_family="AluJo"
     repeat_region   complement(8107. .8411)
                     /rpt_family="AluY"
     repeat_region   9329. .9628
                     /rpt_family="AluSp"
     repeat_region   complement(9687. .9990)
                     /rpt_family="AluSc"
     repeat_region   complement(11137. .11423)
                     /rpt_family="AluSq"
     repeat_region   complement(11476. .11566)
                     /rpt_family="AluSg/x"
     repeat_region   complement(11882. .12185)
                     /rpt_family="MER46C"
     repeat_region   13075. .13377
                     /rpt_family="AluSq"
     repeat_region   13378. .13686
                     /rpt_family="AluSx"
     repeat_region   13808. .13903
                     /rpt_family="L2"
     repeat_region   complement(14066. .14365)
                     /rpt_family="AluSq"
     repeat_region   14652. .14782
                     /rpt_family="L2"
     repeat_region   complement(14951. .15260)
                     /rpt_family="AluSx"
     repeat_region   complement(15772. .16062)
```

Query Match 2.2%; Score 29; DB 9; Length 178217;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||||
Db     165508 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 165536
```

```
RESULT 27
AC090496/c
LOCUS       AC090496               180338 bp    DNA     linear   ROD 20-JUN-2003
DEFINITION  Mus musculus clone rp23-469n6 map 14 strain C57BL/6J, complete
            sequence.
ACCESSION   AC090496
VERSION     AC090496.28  GI:32129354
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 180338)
  AUTHORS   Wu,H., Hu,P., Yang,L., Prescott,A. and Roe,B.A.
  TITLE     Mus musculus BAC Clone rp23-469n6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 180338)
  AUTHORS   Wu,H., Hu,P., Yang,L., Prescott,A. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2001) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 180338)
  AUTHORS   Wu,H., Hu,P., Yang,L., Prescott,A. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 180338)
  AUTHORS   Wu,H., Hu,P., Yang,L., Prescott,A. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Jun 20, 2003 this sequence version replaced gi:31880147.
            ~~~~~~~~~~~~~~ Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ~~~~~~~~~~~~~~~~~~~~~~.
FEATURES             Location/Qualifiers
     source          1. .180338
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /map="14"
                     /clone="rp23-469n6"
                     /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN
```

Query Match 2.2%; Score 29; DB 10; Length 180338;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        812 CCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||
Db      73352 CCTTGGCTGCAAGGGAGTCTGGGAAATGT 73324
```

```
RESULT 28
BC060953
LOCUS       BC060953               723 bp    mRNA    linear   ROD 25-NOV-2003
DEFINITION  Mus musculus cDNA clone MGC:74041 IMAGE:6774045, complete cds.
ACCESSION   BC060953
VERSION     BC060953.1  GI:38511618
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
```

```
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 723)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jonathan Kuo, NIMH
            cDNA Library Preparation: Michael Brownstein /  Ted Usdin
            Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 53 Row: h Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 27734083.
FEATURES             Location/Qualifiers
     source          1. .723
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="MGC:74041 IMAGE:6774045"
                     /tissue_type="Testis, mouse"
                     /clone_lib="NIH_MGC_169"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     CDS             21. .563
                     /codon_start=1
                     /product="Unknown (protein for MGC:74041)"
                     /protein_id="AAH60953.1"
                     /db_xref="GI:38511619"
                     /translation="MARVAKKVHWSRAATAVAATAKAKRSKLKKTAAKRSKLKKTAAK
                     RFKLNKKRNPRSKLPKRSHHSLIHSFIRSRSCGCCHCCCHCCCLHSRPSYRKSTFKIT
                     KKGEQSLRRRIRRRIKRATELRLMQSHLERSQLKLIELEPSEITVALFSHENVNVSEP
                     QEVPPCLDSDPFPNVDLASF"
ORIGIN

  Query Match             2.1%;  Score 28;  DB 10;  Length 723;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-05;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 CAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db        686 CAGATGTCAAAAAAAAAAAAAAAAAAAA 713


RESULT 29
AC004104
LOCUS       AC004104               136037 bp    DNA     linear   PRI 13-JUN-1998
DEFINITION  Homo sapiens 12q24 PAC RPCI5-942N13 (Roswell Park Cancer Institute
            Human PAC library) complete sequence.
ACCESSION   AC004104
VERSION     AC004104.1  GI:3219327
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 136037)
  AUTHORS   Muzny,D., Arenson,A.D., Adams,C., Bunac,C., Carvelli,K., Chang,J.,
            Chacko,J., Chen,J., Ding,Y., Dugan,S., Durbin,J., Forcum,J.,
            Ganesh,R., Garcia,C., Goodman,M., Gorrell,J.H., Haywood,M.,
            Hernandez,J., Jackson,L., Jin,S., Kampal,R., Karpathy,S., Kovar,C.,
            Lau,S., Leal,B., Lee,E., Li,Y., Lichtarge,O., Liu,W., Logan,O.,
            Lu,J., Ly,T., Marondel,I., Martinez,C., Merscher,S., Montgomery,K.,
            Oswal,G., Perez,L., Rashid,N.D., Renault,B., Rowland,K., Savage,L.,
            Scherer,S.E., Shen,H., Simon,M., Stovall,K., Timms,K.M., Todd,J.,
            Vo,Q., Williamson,A., Worley,K.C., Yu,W., Kucherlapati,R.,
            Nelson,D. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 136037)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 136037)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 13, 1998 this sequence version replaced gi:3108008.
            Sequencing is completed to a minimum standard of double strand
            coverage with a minimum of 2 clones and 2 reads with no ambiguities
            or 2 chemistries with a minimum of 2 clones and 3 reads with no
            ambiguities. If the sequence quality does not meet this standard,
            it will be indicated in the annotation.

            The repeat regions shown were identified using RepeatMasker by
            Adrian Smit.

            Sequence similarities were identified using Powerblast by Jinghui
            Zhang.

            Exon/Intron boundaries of identified genes were chosen if there
            were canonical splice junctions that maintained sequence continuity
            across the splice junctions.
FEATURES             Location/Qualifiers
     source          1. .136037
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q24"
                     /clone="RPCI5-942N13"
                     /clone_lib="Roswell Park Cancer Institute Human PAC
                     library"
     repeat_region   1. .113
                     /rpt_family="L1PA11"
     repeat_region   complement(353. .498)
                     /rpt_family="MIR"
     repeat_region   499. .805
                     /rpt_family="AluSp"
     repeat_region   complement(806. .826)
                     /rpt_family="AT_rich"
     repeat_region   complement(905. .995)
                     /rpt_family="(GAAAA)n"
     repeat_region   complement(1848. .2156)
                     /rpt_family="AluSx"
     repeat_region   complement(2299. .2519)
                     /rpt_family="MIR"
     repeat_region   complement(2641. .2806)
                     /rpt_family="MIR"
     repeat_region   2813. .2994
                     /rpt_family="MER53"
     repeat_region   complement(3189. .3307)
                     /rpt_family="MIR"
     repeat_region   3312. .3522
                     /rpt_family="MER20"
     repeat_region   complement(3538. .3656)
                     /rpt_family="MIR"
     repeat_region   complement(4089. .4222)
                     /rpt_family="LTR16C"
```

```
repeat_region   4234. .4764
                /rpt_family="L1PA16"
repeat_region   complement(4786. .4877)
                /rpt_family="LTR16C"
repeat_region   complement(5068. .5103)
                /rpt_family="L2"
repeat_region   complement(5069. .5121)
                /rpt_family="MIR"
repeat_region   5481. .5514
                /rpt_family="AT_rich"
repeat_region   6376. .6597
                /rpt_family="MIR"
repeat_region   7116. .7179
                /rpt_family="MIR"
repeat_region   complement(7197. .7285)
                /rpt_family="L2"
repeat_region   7709. .7851
                /rpt_family="MIR"
repeat_region   complement(7876. .8303)
                /rpt_family="L2"
repeat_region   8308. .8386
                /rpt_family="MIR"
repeat_region   complement(8724. .8818)
                /rpt_family="L1MB6"
repeat_region   8911. .9237
                /rpt_family="AluJo"
repeat_region   complement(9292. .9586)
                /rpt_family="AluSx"
repeat_region   9912. .9980
                /rpt_family="MIR"
repeat_region   10103. .10156
                /rpt_family="L2"
repeat_region   complement(10161. .10506)
                /rpt_family="MLT1B"
repeat_region   complement(10693. .10846)
                /rpt_family="MIR"
repeat_region   11030. .11189
                /rpt_family="L2"
repeat_region   complement(11207. .11503)
                /rpt_family="AluJo"
repeat_region   11557. .11865
                /rpt_family="AluJb"
repeat_region   complement(12046. .12642)
                /rpt_family="L1MB6"
repeat_region   12676. .12962
                /rpt_family="AluJo"
repeat_region   complement(13150. .13192)
                /rpt_family="(TGAA)n"
repeat_region   complement(13193. .13295)
                /rpt_family="L2"
repeat_region   13300. .13582
                /rpt_family="MER7B"
repeat_region   13587. .13866
                /rpt_family="AluSx"
repeat_region   13867. .14110
                /rpt_family="MER7B"
repeat_region   complement(14121. .14228)
                /rpt_family="L1PA14"
repeat_region   14233. .14618
                /rpt_family="MER7B"
repeat_region   complement(14845. .14891)
                /rpt_family="(GAAAA)n"
repeat_region   complement(15165. .15872)
                /rpt_family="TIGGER1"
repeat_region   complement(16050. .16351)
                /rpt_family="AluSx"
repeat_region   complement(16688. .16724)
                /rpt_family="L1PA15"
repeat_region   complement(16947. .17820)
                /rpt_family="L1PA16"
repeat_region   17842. .18169
                /rpt_family="MER7B"
repeat_region   complement(18193. .18536)
                /rpt_family="L2"
repeat_region   complement(18619. .18653)
                /rpt_family="AT_rich"
repeat_region   complement(18654. .18955)
                /rpt_family="AluSx"
repeat_region   complement(18956. .19180)
                /rpt_family="L2"
repeat_region   complement(19502. .20558)
                /rpt_family="L1MC3"
repeat_region   20563. .20675
                /rpt_family="(TA)n"
repeat_region   complement(20693. .20873)
                /rpt_family="L1MD3"
repeat_region   20994. .21182
                /rpt_family="AluJo"
repeat_region   complement(21190. .21321)
                /rpt_family="MIR"
repeat_region   21514. .21604
                /rpt_family="MIR"
repeat_region   21692. .21751
                /rpt_family="MIR"
repeat_region   21749. .21981
                /rpt_family="MIR"
repeat_region   complement(22820. .22879)
                /rpt_family="(TGGA)n"
repeat_region   complement(22891. .22977)
                /rpt_family="MIR"
repeat_region   23027. .23114
                /rpt_family="MER58B"
repeat_region   23169. .23307
                /rpt_family="MER58A"
repeat_region   complement(23539. .23573)
                /rpt_family="(CA)n"
repeat_region   23573. .23612
                /rpt_family="(GA)n"
repeat_region   complement(23695. .23996)
                /rpt_family="AluJo"
repeat_region   complement(24725. .24939)
                /rpt_family="L2"
repeat_region   25378. .25713
                /rpt_family="LTR16C"

  Query Match             2.1%;  Score 28;  DB 9;  Length 136037;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        813 CTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||
Db      65773 CTTGGCTGCAAGGGAGTCTGGGAAATGT 65800


RESULT 30
AC024316
LOCUS       AC024316              137648 bp    DNA     linear   HTG 26-SEP-2000
DEFINITION  Homo sapiens clone RP11-2608, WORKING DRAFT SEQUENCE, 17 unordered
            pieces.
ACCESSION   AC024316
VERSION     AC024316.3  GI:10305255
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 137648)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-2608
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 137648)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
```

```
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 26, 2000 this sequence version replaced gi:10198462.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L4672
                Center clone name: 26_O_8
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 128694 bases at least Q40
                Consensus quality: 133079 bases at least Q30
                Consensus quality: 135006 bases at least Q20
                Insert size: 137000; agarose-fp
                Insert size: 136048; sum-of-contigs
                Quality coverage: 4.7 in Q20 bases; agarose-fp
                Quality coverage: 4.8 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3290: contig of 3290 bp in length
            *     3291     3390: gap of 100 bp
            *     3391     5203: contig of 1813 bp in length
            *     5204     5303: gap of 100 bp
            *     5304     7364: contig of 2061 bp in length
            *     7365     7464: gap of 100 bp
            *     7465    10939: contig of 3475 bp in length
            *    10940    11039: gap of 100 bp
            *    11040    15415: contig of 4376 bp in length
            *    15416    15515: gap of 100 bp
            *    15516    20495: contig of 4980 bp in length
            *    20496    20595: gap of 100 bp
            *    20596    25962: contig of 5367 bp in length
            *    25963    26062: gap of 100 bp
            *    26063    30774: contig of 4712 bp in length
            *    30775    30874: gap of 100 bp
            *    30875    36005: contig of 5131 bp in length
            *    36006    36105: gap of 100 bp
            *    36106    40573: contig of 4468 bp in length
            *    40574    40673: gap of 100 bp
            *    40674    47316: contig of 6643 bp in length
            *    47317    47416: gap of 100 bp
            *    47417    57551: contig of 10135 bp in length
            *    57552    57651: gap of 100 bp
            *    57652    67679: contig of 10028 bp in length
            *    67680    67779: gap of 100 bp
            *    67780    78121: contig of 10342 bp in length
            *    78122    78221: gap of 100 bp
            *    78222   105075: contig of 26854 bp in length
            *   105076   105175: gap of 100 bp
            *   105176   124687: contig of 19512 bp in length
            *   124688   124787: gap of 100 bp
            *   124788   137648: contig of 12861 bp in length.
FEATURES             Location/Qualifiers
     source          1. .137648
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-260B"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .3290
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    3391. .5203
                     /note="assembly_fragment"
     misc_feature    5304. .7364
                     /note="assembly_fragment"
     misc_feature    7465. .10939
                     /note="assembly_fragment"
     misc_feature    11040. .15415
                     /note="assembly_fragment"
     misc_feature    15516. .20495
                     /note="assembly_fragment"
     misc_feature    20596. .25962
                     /note="assembly_fragment"
     misc_feature    26063. .30774
                     /note="assembly_fragment"
     misc_feature    30875. .36005
                     /note="assembly_fragment"
     misc_feature    36106. .40573
                     /note="assembly_fragment"
     misc_feature    40674. .47316
                     /note="assembly_fragment"
     misc_feature    47417. .57551
                     /note="assembly_fragment"
     misc_feature    57652. .67679
                     /note="assembly_fragment"
     misc_feature    67780. .78121
                     /note="assembly_fragment"
     misc_feature    78222. .105075
                     /note="assembly_fragment"
     misc_feature    105176. .124687
                     /note="assembly_fragment"
     misc_feature    124788. .137648
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             2.1%;  Score 28;  DB 2;  Length 137648;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 CAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db      19515 CAGATGTCAAAAAAAAAAAAAAAAAAAA 19542


RESULT 31
AC051635/c
LOCUS       AC051635              158608 bp    DNA     linear   PRI 27-AUG-2002
DEFINITION  Homo sapiens chromosome 18, clone RP11-540A4, complete sequence.
ACCESSION   AC051635
VERSION     AC051635.7  GI:17298634
```

KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 158608)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-540A4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 158608)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 158608)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 158608)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 5 (bases 1 to 158608)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Dec 4, 2001 this sequence version replaced gi:16756307.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L5261
Center clone name: 540_A_4
----------.
FEATURES Location/Qualifiers
source 1. .158608
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-540A4"
/clone_lib="RPCI-11 Human Male BAC"
repeat_region 8. .106
/rpt_family="AluSg/x"
unsure 15. .27
/note="<30 qual SNGL region"
unsure 75. .77
/note="<30 qual SNGL region"
unsure 106. .109
/note="<30 qual SNGL region"
repeat_region 112. .228
/rpt_family="L1MB2"
repeat_region complement(563. .632)
/rpt_family="MIR"

```
     unsure          1053. .1057
                     /note="<30 qual SNGL region"
     repeat_region   complement(1123. .1352)
                     /rpt_family="MIR"
     repeat_region   1624. .1702
                     /rpt_family="L2"
     repeat_region   complement(1853. .2218)
                     /rpt_family="L2"
     repeat_region   3091. .3358
                     /rpt_family="L1MC5"
     repeat_region   3357. .3850
                     /rpt_family="L1MC1"
     repeat_region   complement(3868. .4006)
                     /rpt_family="LTR7"
     repeat_region   4011. .4167
                     /rpt_family="MIR3"
     repeat_region   complement(4235. .4560)
                     /rpt_family="LTR16C"
     repeat_region   5345. .5372
                     /rpt_family="(TTTG)n"
     repeat_region   5957. .5984
                     /rpt_family="(TTTG)n"
     repeat_region   6132. .6246
                     /rpt_family="MIR3"
     repeat_region   6598. .6718
                     /rpt_family="L1MEd"
     repeat_region   6731. .6769
                     /rpt_family="(CA)n"
     repeat_region   6940. .7148
                     /rpt_family="L1MEd"
     repeat_region   7153. .7466
                     /rpt_family="L1MB8"
     repeat_region   7510. .8250
                     /rpt_family="L1MEd"
     repeat_region   8279. .8516
                     /rpt_family="L1MEd"
     repeat_region   8612. .8827
                     /rpt_family="AluJb"
     repeat_region   9027. .9126
                     /rpt_family="L1MEd"
     repeat_region   9127. .9435
                     /rpt_family="AluY"
     repeat_region   9436. .9486
                     /rpt_family="L1MEd"
     repeat_region   9487. .9805
                     /rpt_family="AluSx"
     repeat_region   9806. .9922
                     /rpt_family="L1MEd"
     repeat_region   complement(9940. .9989)
                     /rpt_family="MLT1I"

  Query Match             2.1%;  Score 28;  DB 9;  Length 158608;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 CAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db      82163 CAGATGTCAAAAAAAAAAAAAAAAAAAA 82136


RESULT 32
AC069032
LOCUS       AC069032              187557 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 18 clone RP11-185D6 map 18, WORKING DRAFT
            SEQUENCE, 44 unordered pieces.
ACCESSION   AC069032
VERSION     AC069032.4  GI:12963008
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187557)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-185D6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 187557)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 187557)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 18, 2001 this sequence version replaced gi:11225387.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L10383
                Center clone name: 185_D_6
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 2% of reads
                Sequencing vector: Plasmid; n/a; 98% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 166116 bases at least Q40
                Consensus quality: 176954 bases at least Q30
                Consensus quality: 180848 bases at least Q20
```

```
    Insert size: 167000; agarose-fp
    Insert size: 183257; sum-of-contigs
    Quality coverage: 3.7 in Q20 bases; agarose-fp
    Quality coverage: 3.4 in Q20 ba.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1294: contig of 1294 bp in length
*     1295     1394: gap of 100 bp
*     1395     2505: contig of 1111 bp in length
*     2506     2605: gap of 100 bp
*     2606     3992: contig of 1387 bp in length
*     3993     4092: gap of 100 bp
*     4093     5527: contig of 1435 bp in length
*     5528     5627: gap of 100 bp
*     5628     6675: contig of 1048 bp in length
*     6676     6775: gap of 100 bp
*     6776     8026: contig of 1251 bp in length
*     8027     8126: gap of 100 bp
*     8127     9162: contig of 1036 bp in length
*     9163     9262: gap of 100 bp
*     9263    10351: contig of 1089 bp in length
*    10352    10451: gap of 100 bp
*    10452    11757: contig of 1306 bp in length
*    11758    11857: gap of 100 bp
*    11858    12878: contig of 1021 bp in length
*    12879    12978: gap of 100 bp
*    12979    14159: contig of 1181 bp in length
*    14160    14259: gap of 100 bp
*    14260    15818: contig of 1559 bp in length
*    15819    15918: gap of 100 bp
*    15919    17145: contig of 1227 bp in length
*    17146    17245: gap of 100 bp
*    17246    18765: contig of 1520 bp in length
*    18766    18865: gap of 100 bp
*    18866    20138: contig of 1273 bp in length
*    20139    20238: gap of 100 bp
*    20239    21879: contig of 1641 bp in length
*    21880    21979: gap of 100 bp
*    21980    23877: contig of 1898 bp in length
*    23878    23977: gap of 100 bp
*    23978    25680: contig of 1703 bp in length
*    25681    25780: gap of 100 bp
*    25781    28099: contig of 2319 bp in length
*    28100    28199: gap of 100 bp
*    28200    30797: contig of 2598 bp in length
*    30798    30897: gap of 100 bp
*    30898    33191: contig of 2294 bp in length
*    33192    33291: gap of 100 bp
*    33292    36520: contig of 3229 bp in length
*    36521    36620: gap of 100 bp
*    36621    38842: contig of 2222 bp in length
*    38843    38942: gap of 100 bp
*    38943    42470: contig of 3528 bp in length
*    42471    42570: gap of 100 bp
*    42571    45581: contig of 3011 bp in length
*    45582    45681: gap of 100 bp
*    45682    48005: contig of 2324 bp in length
*    48006    48105: gap of 100 bp
*    48106    51907: contig of 3802 bp in length
*    51908    52007: gap of 100 bp
*    52008    55311: contig of 3304 bp in length
*    55312    55411: gap of 100 bp
*    55412    59353: contig of 3942 bp in length
*    59354    59453: gap of 100 bp
*    59454    63506: contig of 4053 bp in length
*    63507    63606: gap of 100 bp
*    63607    68052: contig of 4446 bp in length
*    68053    68152: gap of 100 bp
*    68153    72726: contig of 4574 bp in length
*    72727    72826: gap of 100 bp
*    72827    78593: contig of 5767 bp in length
*    78594    78693: gap of 100 bp
*    78694    84391: contig of 5698 bp in length
*    84392    84491: gap of 100 bp
*    84492    92490: contig of 7999 bp in length
*    92491    92590: gap of 100 bp
*    92591    99730: contig of 7140 bp in length
*    99731    99830: gap of 100 bp
*    99831   107602: contig of 7772 bp in length
*   107603   107702: gap of 100 bp
*   107703   116403: contig of 8701 bp in length
*   116404   116503: gap of 100 bp
*   116504   126259: contig of 9756 bp in length
*   126260   126359: gap of 100 bp
*   126360   136166: contig of 9807 bp in length
*   136167   136266: gap of 100 bp
*   136267   145578: contig of 9312 bp in length
*   145579   145678: gap of 100 bp
*   145679   157186: contig of 11508 bp in length
*   157187   157286: gap of 100 bp
*   157287   170785: contig of 13499 bp in length
*   170786   170885: gap of 100 bp
*   170886   187557: contig of 16672 bp in length.
FEATURES             Location/Qualifiers
     source          1. .187557
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-185D6"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1294
                     /note="assembly_fragment"
     misc_feature    1395. .2505
                     /note="assembly_fragment"
     misc_feature    2606. .3992
                     /note="assembly_fragment"
     misc_feature    4093. .5527
                     /note="assembly_fragment"

  Query Match             2.1%;  Score 28;  DB 2;  Length 187557;
  Best Local Similarity   100.0%;  Pred. No. 0.00014;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 CAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db      28971 CAGATGTCAAAAAAAAAAAAAAAAAAAA 28998


RESULT 33
AC068678/c
LOCUS       AC068678               193978 bp    DNA     linear   HTG 31-OCT-2001
DEFINITION  Homo sapiens chromosome 18 clone RP11-104N14 map 18, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC068678
VERSION     AC068678.3  GI:16506881
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193978)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-104N14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 193978)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
```

```
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Oct 28, 2001 this sequence version replaced gi:13249473.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1295
                Center clone name: 104_N_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 4% of reads
                Sequencing vector: Plasmid; n/a; 96% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 191018 bases at least Q40
                Consensus quality: 192308 bases at least Q30
                Consensus quality: 192875 bases at least Q20
                Insert size: 190000; agarose-fp
                Insert size: 193578; sum-of-contigs
                Quality coverage: 7.3 in Q20 bases; agarose-fp
                Quality coverage: 7.1 in Q20 ba.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    59978: contig of 59978 bp in length
            *    59979    60078: gap of 100 bp
            *    60079    83937: contig of 23859 bp in length
            *    83938    84037: gap of 100 bp
            *    84038   113621: contig of 29584 bp in length
            *   113622   113721: gap of 100 bp
            *   113722   147302: contig of 33581 bp in length
            *   147303   147402: gap of 100 bp
            *   147403   193978: contig of 46576 bp in length.
FEATURES            Location/Qualifiers
     source         1. .193978
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="18"
                    /map="18"
                    /clone="RP11-104N14"
                    /clone_lib="RPCI-11 Human Male BAC"
     misc_feature   1. .59978
                    /note="assembly_fragment
                    clone_end:SP6
                    vector_side:left"
     misc_feature   60079. .83937
                    /note="assembly_fragment"
     misc_feature   84038. .113621
                    /note="assembly_fragment"
     misc_feature   113722. .147302
                    /note="assembly_fragment"
     misc_feature   147403. .193978
                    /note="assembly_fragment
                    clone_end:T7
                    vector_side:right"
ORIGIN

  Query Match             2.1%;  Score 28;  DB 2;  Length 193978;
  Best Local Similarity   100.0%;  Pred. No. 0.00014;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 CAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db      20268 CAGATGTCAAAAAAAAAAAAAAAAAAAA 20241


RESULT 34
AC011814
LOCUS       AC011814               209462 bp    DNA     linear   HTG 26-MAY-2000
DEFINITION  Homo sapiens chromosome 18 clone RP11-153D16 map 18, WORKING DRAFT
            SEQUENCE, 40 unordered pieces.
ACCESSION   AC011814
VERSION     AC011814.3  GI:8072439
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 209462)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-153D16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209462)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 25, 2000 this sequence version replaced gi:7321509.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1307
                Center clone name: 153_D_16
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 97% of reads
                Sequencing vector: Plasmid; n/a; %-0.f%% of reads
                3.29835082458771Chemistry: Dye-primer-amersham; 8% of reads
```

```
    Chemistry: Dye-terminator Big Dye; 92% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 172516 bases at least Q40
    Consensus quality: 188069 bases at least Q30
    Consensus quality: 195901 bases at least Q20
    Insert size: 175000; agarose-fp
    Insert size: 205562; sum-of-contigs
    Quality covera.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1353: contig of 1353 bp in length
*     1354     1453: gap of 100 bp
*     1454     2778: contig of 1325 bp in length
*     2779     2878: gap of 100 bp
*     2879     4211: contig of 1333 bp in length
*     4212     4311: gap of 100 bp
*     4312     5622: contig of 1311 bp in length
*     5623     5722: gap of 100 bp
*     5723     7256: contig of 1534 bp in length
*     7257     7356: gap of 100 bp
*     7357     8828: contig of 1472 bp in length
*     8829     8928: gap of 100 bp
*     8929    10225: contig of 1297 bp in length
*    10226    10325: gap of 100 bp
*    10326    11376: contig of 1051 bp in length
*    11377    11476: gap of 100 bp
*    11477    12643: contig of 1167 bp in length
*    12644    12743: gap of 100 bp
*    12744    13857: contig of 1114 bp in length
*    13858    13957: gap of 100 bp
*    13958    15355: contig of 1398 bp in length
*    15356    15455: gap of 100 bp
*    15456    17127: contig of 1672 bp in length
*    17128    17227: gap of 100 bp
*    17228    19117: contig of 1890 bp in length
*    19118    19217: gap of 100 bp
*    19218    20984: contig of 1767 bp in length
*    20985    21084: gap of 100 bp
*    21085    23101: contig of 2017 bp in length
*    23102    23201: gap of 100 bp
*    23202    25331: contig of 2130 bp in length
*    25332    25431: gap of 100 bp
*    25432    27214: contig of 1783 bp in length
*    27215    27314: gap of 100 bp
*    27315    29160: contig of 1846 bp in length
*    29161    29260: gap of 100 bp
*    29261    32570: contig of 3310 bp in length
*    32571    32670: gap of 100 bp
*    32671    34880: contig of 2210 bp in length
*    34881    34980: gap of 100 bp
*    34981    38325: contig of 3345 bp in length
*    38326    38425: gap of 100 bp
*    38426    42527: contig of 4102 bp in length
*    42528    42627: gap of 100 bp
*    42628    46099: contig of 3472 bp in length
*    46100    46199: gap of 100 bp
*    46200    49618: contig of 3419 bp in length
*    49619    49718: gap of 100 bp
*    49719    53310: contig of 3592 bp in length
*    53311    53410: gap of 100 bp
*    53411    56772: contig of 3362 bp in length
*    56773    56872: gap of 100 bp
*    56873    60358: contig of 3486 bp in length
*    60359    60458: gap of 100 bp
*    60459    66617: contig of 6159 bp in length
*    66618    66717: gap of 100 bp
*    66718    75403: contig of 8686 bp in length
*    75404    75503: gap of 100 bp
*    75504    83115: contig of 7612 bp in length
*    83116    83215: gap of 100 bp
*    83216    91822: contig of 8607 bp in length
*    91823    91922: gap of 100 bp
*    91923   101265: contig of 9343 bp in length
*   101266   101365: gap of 100 bp
*   101366   111447: contig of 10082 bp in length
*   111448   111547: gap of 100 bp
*   111548   121183: contig of 9636 bp in length
*   121184   121283: gap of 100 bp
*   121284   133655: contig of 12372 bp in length
*   133656   133755: gap of 100 bp
*   133756   147893: contig of 14138 bp in length
*   147894   147993: gap of 100 bp
*   147994   160302: contig of 12309 bp in length
*   160303   160402: gap of 100 bp
*   160403   174936: contig of 14534 bp in length
*   174937   175036: gap of 100 bp
*   175037   189972: contig of 14936 bp in length
*   189973   190072: gap of 100 bp
*   190073   209462: contig of 19390 bp in length.
FEATURES             Location/Qualifiers
     source          1. .209462
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-153D16"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1353
                     /note="assembly_fragment"
     misc_feature    1454. .2778
                     /note="assembly_fragment"
     misc_feature    2879. .4211
                     /note="assembly_fragment"
     misc_feature    4312. .5622
                     /note="assembly_fragment"
     misc_feature    5723. .7256
                     /note="assembly_fragment"
     misc_feature    7357. .8828
                     /note="assembly_fragment"
     misc_feature    8929. .10225
                     /note="assembly_fragment"
     misc_feature    10326. .11376
                     /note="assembly_fragment"
     misc_feature    11477. .12643
                     /note="assembly_fragment"
     misc_feature    12744. .13857
                     /note="assembly_fragment"
     misc_feature    13958. .15355
                     /note="assembly_fragment"
     misc_feature    15456. .17127
                     /note="assembly_fragment"
     misc_feature    17228. .19117
                     /note="assembly_fragment"
     misc_feature    19218. .20984
                     /note="assembly_fragment"
     misc_feature    21085. .23101
                     /note="assembly_fragment"
     misc_feature    23202. .25331
                     /note="assembly_fragment"
     misc_feature    25432. .27214
                     /note="assembly_fragment"
     misc_feature    27315. .29160
                     /note="assembly_fragment"
     misc_feature    29261. .32570
                     /note="assembly_fragment"
     misc_feature    32671. .34880
                     /note="assembly_fragment"
     misc_feature    34981. .38325
                     /note="assembly_fragment"
```

```
     misc_feature    38426. .42527
                     /note="assembly_fragment"
     misc_feature    42628. .46099
                     /note="assembly_fragment"
     misc_feature    46200. .49618
                     /note="assembly_fragment"

  Query Match              2.1%;  Score 28;  DB 2;  Length 209462;
  Best Local Similarity   100.0%;  Pred. No. 0.00014;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 CAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db      54971 CAGATGTCAAAAAAAAAAAAAAAAAAAA 54998


RESULT 35
BC060625
LOCUS       BC060625               4150 bp    mRNA    linear   ROD 12-NOV-2003
DEFINITION  Mus musculus Tnf receptor-associated factor 2, mRNA (cDNA clone
            IMAGE:6831932), partial cds.
ACCESSION   BC060625
VERSION     BC060625.1  GI:38197727
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4150)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 4150)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
            Thomas L. Casavant.
            Web site: http://genome.uiowa.edu
            Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
            Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
            Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
            Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
            Casavant,T., Soares,M.B.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series:  Plate:  Row:  Column: 0
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .4150
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6831932"
                     /tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc"
                     /clone_lib="NIH_BMAP_FY0"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
ORIGIN

  Query Match              2.0%;  Score 27;  DB 10;  Length 4150;
  Best Local Similarity   100.0%;  Pred. No. 0.00041;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db       4111 AGATGTCAAAAAAAAAAAAAAAAAAAA 4137


RESULT 36
HSM803417
LOCUS       HSM803417              4685 bp    mRNA    linear   PRI 13-MAY-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp313H168 (from clone DKFZp313H168).
ACCESSION   AL832110
VERSION     AL832110.1  GI:21732653
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4685)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp313H168) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .4685
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313H168"
                     /tissue_type="human bone marrow"
                     /clone_lib="313 (synonym: hlcc2). Vector pTriplEx2; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
     polyA_signal    4647. .4652
     polyA_site      4662
ORIGIN

  Query Match              2.0%;  Score 27;  DB 9;  Length 4685;
  Best Local Similarity   100.0%;  Pred. No. 0.00041;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||
Db       4656 AGATGTCAAAAAAAAAAAAAAAAAAA 4682


RESULT 37
AC022522/c
LOCUS       AC022522               88643 bp    DNA     linear   PLN 12-SEP-2000
DEFINITION  Arabidopsis thaliana chromosome I BAC T28K15 genomic sequence,
            complete sequence.
ACCESSION   AC022522
VERSION     AC022522.2  GI:6970637
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 88643)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Miranda,M., Brooks,S.,
            Buehler,E., Chao,Q., Chin,C., Chiou,J., Choi,E., Gonzalez,A.,
            Howng,B., Johnson-Hopson,C., Khan,S., Kim,C., Koo,T., Lee,J.M.,
            Lenz,C., Liu,A., Liu,S., Mukharsky,N., Pham,P., Sakano,H.,
            Shinn,P., Toriumi,M., Vaysberg,M., Yu,G., Ecker,J., Theologis,A.
            and Davis,R.W.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 88643)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,
            Chin,C., Chiou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C.,
            Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
            Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,
            Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
            Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2000) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   3  (bases 1 to 88643)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,
            Chin,C., Chiou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C.,
            Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
            Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,
            Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
            Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-2000) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   4  (bases 1 to 88643)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
            and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Stanford Genome Technology Center, Stanford
            University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT     On Feb 14, 2000 this sequence version replaced gi:6910565.
            Bases 57,901-88,643 of clone T28K15 overlap with bases 1-30,740 of
            IGF BAC clone F12F1 (AC002131).
             e-mail for correspondence: arab@sequence.stanford.edu
            Genes with similarity to proteins in the databases are named
            'putative', '-like' or 'similar to'.  Genes that have EST
            similarity but no significant protein similarity are described as
            'unknown proteins'.  Genes that are annotated based only on gene
            prediction software are described as 'hypothetical proteins'.   The
            gene prediction programs used to predict genes include:  Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
            http://genes.mit.edu/GENSCAN.), FEXA (Victor Solovyev,
            http://genomic.sanger.ac.uk/gf/gf.shtml),  and NetPlantGene (S.M.
            Hebsgaard, et al., CBS, Technical University of
            Denmark,http://www.cbs.dtu.dk/NetPlantGene.html).
FEATURES            Location/Qualifiers
     source         1. .88643
                    /organism="Arabidopsis thaliana"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:3702"
                    /chromosome="I"
                    /map="near NCC1"
                    /clone="T28K15"
     misc_feature   1. .1849
                    /note="overlaps with BAC F5O11 from 1-1849"
     gene           8039. .9467
                    /gene="T28K15.1"
     CDS            join(8039. .8198,8253. .8566,8712. .9467)
                    /gene="T28K15.1"
                    /note="Hypothetical protein"
                    /codon_start=1
                    /protein_id="AAG12568.1"
                    /db_xref="GI:10086508"
                    /translation="MNSFSHVPPGFRFHPTDEELVDYYLRKKVASKRIEIDFIKDIDL
                    YKIEPWDLQGKILKVFFYSFVEIELCKIGHEEQSDWYFFSHKDKKYPTGTRTNRATKA
                    GFWKATGRDKAIYLRHSLIGMRKTLVFYKGRAPNGQKSDWIMHEYRLETDENGTPQEE
                    GWVVCRVFKKRLAAVRRMGDYDSSPSHWYDDQLSFMASELETNGQRRILPNHHQQQQH
                    EHQQHMPYGLNASAYALNNPNLQCKQELELHYNHLVQRNHLLDESHLSFLQLPQLESP
                    KIQQDNSNCNSLPYGTSNIDNNSSHNANLQQSNIAHEEQLNQGNQNFSSLYMNSGNEQ
                    VMDQVTDWRVLDKFVASQLSNEEAATASASIQNNAKDTSNAEYQVDEEKDPKRASDMG
                    EEYTASTSSSCQIDLWK"
     gene           complement(11256. .13208)
                    /gene="T28K15.2"
     CDS            complement(join(11256. .11663,11757. .11791,12355. .12461,
                    12642. .12671,12775. .12942,13075. .13208))
                    /gene="T28K15.2"
                    /codon_start=1
                    /product="Hypothetical protein"
                    /protein_id="AAG12570.1"
                    /db_xref="GI:10086510"
                    /translation="MKSLDISRSSSSVSRSPYHFQRYLLRRLQLSSRSNLEIKDSSNT
                    SAESNTWKRILSAAMAAAVIASSSGVPAMAELNRFEADTRGEFGIGSAAQYGSADLSK
                    TVHSNENFRRANFTSADMRESDFSGSTFNGAYLEKAVAYKANFSGADLSDTLMDRMVL
                    NEANLTNAVLVRSVLTRSDLGGAKIEGADFSDAVIDLLQKQVTTTHHYIYPSFRSTIK
                    KYFTNGFHNVLKALCKYATGTNPLTGVDTRKSLGCGNSRRNAYGSPSSPLLSAPPQRL
                    LGRDGFCDEKTGLCDVK"
     gene           complement(15068. .18826)
                    /gene="T28K15.3"
     CDS            complement(join(15068. .15286,15376. .15463,15549. .15961,
                    16068. .16685,17191. .17435,18087. .18095,18424. .18826))
                    /gene="T28K15.3"
                    /codon_start=1
                    /product="beta-fructosidase"
                    /protein_id="AAG12569.1"
                    /db_xref="GI:10086509"
                    /translation="MASSDALLPISAREEEPLCPYTRLPMADPNQETHGPRRRRPFKG
                    LLAVSFGLLFIAFYVALIATHDGSRSNDEGIDETETITSRARLAGVSEKRNDGLWKLS
                    GDRNTPAFEWNNSMLSWQRTAFHFQPEQNWMNDPNGPLFYKGWYHFFYQYNPNAAVWG
                    DIVWGHAVSRDLIHWVHLPIAMVADQWYDSNGVWTGSATFLPDGSIVMLYTGSTDKAV
                    QVQNLAYPEDPNDPLLLKWVKFPGNPVLVPPPGILPKDFRDPTTAWKTSEGKWRITIG
                    SKLNKTGISLVYDTIDPKTYEKLDTLLHRVPNTGMWECVDFYPVSKTAGNGLDTSVNG
                    PDVKHIVKASMDDTRFDHYAVGTYFDSNGTWIPDDPTIDVGMTASLRYDYGKFYASKS
                    FYDQNKGRRVLWSWIGESDSEASDVQKGWSSLQGIPRTVVLDTKTGKNLVQWPVEEIK
                    SLRLSSKQFDLEVGPGSVVPVDVGSAAQLDIEAEFEINKESLDKIIGNASVVAEAEEF
                    SCEKSGGSTVRGALGPFGFSVLATESLSEQTPVYFYVAKGKDSELKTFFCTDTSRSSV
                    ANDVVKPIYGSVVPVLKGEKLTMRILVDHSIVEAFGQGGRTCITSRVYPTTAIYGAAK
                    LFLFNNALDATVTASFTVWQMNSAFIHPYSDEAVRALSRT"
     gene           complement(21817. .24475)
                    /gene="T28K15.4"
     CDS            complement(join(21817. .21963,22048. .22186,22271. .22446,
                    22520. .22589,22684. .22748,22824. .22904,22994. .23098,
                    23192. .23275,23370. .23445,23708. .23786,24190. .24331,
                    24422. .24475))
                    /gene="T28K15.4"
                    /note="Similar to transaldolase"
                    /codon_start=1
                    /protein_id="AAG12571.1"
```

```
                     /db_xref="GI:10086511"
                     /translation="MPLSLQSPPCATLSASIRKGRWKTLAAGFSVTPLPAVNFSLRRS
                     IPRILASASSSSSPASSSLEAGENNELNAVSAFSEIVPDTVVFDDFERFPPTAATVSS
                     ALLLGICGLPDTIFRNAVDMALADSSCAGLETTESRLSCFFNKAIVNVGGDLVKLVPG
                     RVSTEVDARLAYDTNGIIRKVHDLLRLYNEIDVPHDRLLFKIPATWQGIEAARLLESE
                     GIQTHMTFVYSFAQAAAASQAGASVIQIFVGRLRDWARNHSGDTEIESAIKSGEDPGL
                     ALVKRSYNYIHKYGYKSKLMAAAVRNKQDLFSLLGVDYVIAPLKVLQSLKDSPAIPDD
                     EKYSFVRKLTPETATHYHFTNKELVKWDQLSLASSMGPASVELLSAGVEGYANQAKRV
                     EELFGKIWPPPNV"
     gene            complement(24845. .27514)
                     /gene="T28K15.5"
     CDS             complement(24845. .27514)
                     /gene="T28K15.5"
                     /note="NBS/LRR disease resistance gene, location of
                     genomic clone gb|AF074916."
                     /codon_start=1
                     /product="resistance to Pseudomonas syringae protein 5"
                     /protein_id="AAG12572.1"
                     /db_xref="GI:10086512"
                     /translation="MGGCFSVSLPCDQVVSQFSQLLCVRGSYIHNLSKNLASLQKAMR
                     MLKARQYDVIRRLETEEFTGRQQRLSQVQVWLTSVLIIQNQFNDLLRSNEVELQRLCL
                     CGFCSKDLKLSYRYGKRVIMMLKEVESLSSQGFFDVVSEATPFADVDEIPFQPTIVGQ
                     EIMLEKAWNRLMEDGSGILGLYGMGGVGKTTLLTKINNKFSKIDDRFDVVIWVVVSRS
                     STVRKIQRDIAEKVGLGGMEWSEKNDNQIAVDIHNVLRRRKFVLLLDDIWEKVNLKAV
                     GVPYPSKDNGCKVAFTTRSRDVCGRMGVDDPMEVSCLQPEESWDLFQMKVGKNTLGSH
                     PDIPGLARKVARKCRGLPLALNVIGEAMACKRTVHEWCHAIDVLTSSAIDFSGMEDEI
                     LHVLKYSYDNLNGELMKSCFLYCSLFPEDYLIDKEGLVDYWISEGFINEKEGRERNIN
                     QGYEIIGTLVRACLLLEEERNKSNVKMHDVVREMALWISSDLGKQKEKCIVRAGVGLR
                     EVPKVKDWNTVRKISLMNNEIEEIFDSHECAALTTLFLQKNDVVKISAEFFRCMPHLV
                     VLDLSENQSLNELPEEISELASLRYFNLSYTCIHQLPVGLWTLKKLIHLNLEHMSSLG
                     SILGISNLWNLRTLGLRDSRLLLDMSLVKELQLLEHLEVITLDISSSLVAEPLLCSQR
                     LVECIKEVDFKYLKEESVRVLTLPTMGNLRKLGIKRCGMREIKIERTTSSSSRNKSPT
                     TPCFSNLSRVFIAKCHGLKDLTWLLFAPNLTFLEVGFSKEVEDIISEEKAEEHSATIV
                     PFRKLETLHLFELRGLKRIYAKALHFPCLKVIHVEKCEKLRKLPLDSKSGIAGEELVI
                     YYGEREWIERVEWEDQATQLRFLPSSRWRWRET"
     gene            complement(28920. .31577)
                     /gene="T28K15.6"
     CDS             complement(28920. .31577)
                     /gene="T28K15.6"
                     /note="Similar to RPS5, location of genomic clone
                     gb|AF074916."
                     /codon_start=1
                     /product="NBS/LRR disease resistance protein"
                     /protein_id="AAG12573.1"
                     /db_xref="GI:10086513"
                     /translation="MGGCVSVSLSCDREVNQFSQWLCVSGSYIQNLSENLASLQKAMG
                     VLNAKRDDVQGRINREEFTGHRRRLAQVQVWLTRIQTIENQFNDLLSTCNAEIQRLCL
                     CGFCSKNVKMSYLYGKRVIVLLREVEGLSSQGVFDIVTEAAPIAEVEELPIQSTIVGQ
                     DSMLDKVWNCLMEDKVWIVGLYGMGGVGKTTLLTQINNKFSKLGGGFDVVIWVVVSKN
                     ATVHKIQKSIGEKLGLVGKNWDEKNKNQRALDIHNVLRRKKFVLLLDDIWEKVELKVI
                     GVPYPSGENGCKVAFTTHSKEVCGRMGVDNPMEISCLDTGNAWDLLKKKVGENTLGSH
                     PDIPQLARKVSEKCCGLPLALNVIGETMSFKRTIQEWRHATEVLTSATDFSGMEDEIL
                     PILKYSYDSLNGEDAKSCFLYCSLFPEDFEIRKEMLIEYWICEGFIKEKQGREKAFNQ
                     GYDILGTLVRSSLLLEGAKDKDVVSMHDMVREMALWIFSDLGKHKERCIVQAGIGLDE
                     LPEVENWRAVKRMSLMNNNFEKILGSPECVELITLFLQNNYKLVDISMEFFRCMPSLA
                     VLDLSENHSLSELPEEISELVSLQYLDLSGTYIERLPHGLHELRKLVHLKLERTRRLE
                     SISGISYLSSLRTLRLRDSKTTLDTGLMKELQLLEHLELITTDISSGLVGELFCYPRV
                     GRCIQHIYIRDHWERPEESVGVLVLPAIHNLCYISIWNCWMWEIMIEKTPWKKNLTNP
                     NFSNLSNVRIEGCDGLKDLTWLLFAPNLINLRVWGCKHLEDIISKEKAASVLEKEILP
                     FQKLECLNLYQLSELKSIYWNALPFQRLRCLDILNNCPKLRKLPLDSKSVVKVEEFVI
                     KYKEKKWIERVEWEDEATQYRFLPTCRLR"

  Query Match             2.0%;  Score 27;  DB 8;  Length 88643;
  Best Local Similarity   100.0%;  Pred. No. 0.00052;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db       6415 AGATGTCAAAAAAAAAAAAAAAAAAAA 6389


RESULT 38
AP001576
```

```
LOCUS       AP001576              105199 bp    DNA     linear   PRI 16-AUG-2001
DEFINITION  Homo sapiens genomic DNA, chromosome 6q25.2, clone:KB1F5.
ACCESSION   AP001576
VERSION     AP001576.3  GI:15208253
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimizu,N. and Asakawa,S.
  TITLE     Homo sapiens DNA chromosome 6 SEQUENCE
  JOURNAL   Published Only in Database (2000)
REFERENCE   2  (bases 1 to 105199)
  AUTHORS   Shimizu,N. and Asakawa,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Nobuyoshi Shimizu, Keio University, School
            of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo
            160-8582, Japan (E-mail:nshimizu@dmb.med.keio.ac.jp,
            Tel:81-3-3351-2370, Fax:81-3-3351-2370)
COMMENT     On Aug 16, 2001 this sequence version replaced gi:8096256.
FEATURES             Location/Qualifiers
     source          1. .105199
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6q25.2"
                     /clone="KB1F5"
                     /cell_line="FLEB 14 - 14"
                     /clone_lib="Keio BAC library"
     repeat_region   166. .474
                     /evidence=not_experimental
                     /rpt_family="AluSx"
     repeat_region   520. .590
                     /evidence=not_experimental
                     /rpt_family="AT_rich"
     repeat_region   complement(912. .1017)
                     /evidence=not_experimental
                     /rpt_family="FLAM_C"
     repeat_region   1199. .1490
                     /evidence=not_experimental
                     /rpt_family="MLT1I"
     repeat_region   1844. .1922
                     /evidence=not_experimental
                     /rpt_family="(TATATG)n"
     repeat_region   2732. .3028
                     /evidence=not_experimental
                     /rpt_family="AluSx"
     repeat_region   complement(3323. .3404)
                     /evidence=not_experimental
                     /rpt_family="L2"
     repeat_region   complement(3608. .3816)
                     /evidence=not_experimental
                     /rpt_family="L2"
     repeat_region   complement(4573. .4720)
                     /evidence=not_experimental
                     /rpt_family="MIR"
     repeat_region   complement(7641. .8105)
                     /evidence=not_experimental
                     /rpt_family="L1PA16"
     repeat_region   8106. .8852
                     /evidence=not_experimental
                     /rpt_family="L1PA16"
     repeat_region   8855. .8884
                     /evidence=not_experimental
                     /rpt_family="AT_rich"
     repeat_region   9163. .9192
                     /evidence=not_experimental
                     /rpt_family="AT_rich"
     repeat_region   9351. .9656
                     /evidence=not_experimental
                     /rpt_family="AluSx"
```

```
repeat_region   complement(10418. .10545)
                /evidence=not_experimental
                /rpt_family="MER58C"
repeat_region   10546. .10585
                /evidence=not_experimental
                /rpt_family="Alu"
repeat_region   10708. .10797
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   complement(10837. .11091)
                /evidence=not_experimental
                /rpt_family="AluSc"
repeat_region   11771. .11813
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   12272. .12292
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   complement(12629. .12706)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   12903. .13167
                /evidence=not_experimental
                /rpt_family="AluJo"
repeat_region   13173. .13453
                /evidence=not_experimental
                /rpt_family="AluSg"
repeat_region   13871. .13920
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   complement(14938. .15193)
                /evidence=not_experimental
                /rpt_family="AluSq"
repeat_region   15379. .15453
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   complement(15501. .15754)
                /evidence=not_experimental
                /rpt_family="LTR37B"
repeat_region   complement(16177. .16924)
                /evidence=not_experimental
                /rpt_family="MER31-internal"
repeat_region   complement(17232. .17377)
                /evidence=not_experimental
                /rpt_family="MER31-internal"
repeat_region   18854. .18955
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   19314. .19624
                /evidence=not_experimental
                /rpt_family="AluSx"
repeat_region   20973. .21071
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(22061. .22369)
                /evidence=not_experimental
                /rpt_family="AluSx"
repeat_region   22428. .22463
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   complement(22469. .22781)
                /evidence=not_experimental
                /rpt_family="AluJo"
repeat_region   complement(22820. .23154)
                /evidence=not_experimental
                /rpt_family="MER7A"
repeat_region   23362. .23397
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   23468. .23503
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   24148. .24175
                /evidence=not_experimental
                /rpt_family="(A)n"
repeat_region   complement(24576. .24879)
                /evidence=not_experimental
                /rpt_family="AluSx"
repeat_region   24980. .25344
                /evidence=not_experimental
                /rpt_family="MLT2E"
repeat_region   complement(25365. .25422)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(25372. .25487)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   25515. .25637
                /evidence=not_experimental
                /rpt_family="MER94"
repeat_region   25822. .25868
                /evidence=not_experimental
                /rpt_family="Alu"
repeat_region   26514. .26818
                /evidence=not_experimental
                /rpt_family="AluYa5"
repeat_region   27318. .27344
                /evidence=not_experimental
                /rpt_family="(TTTG)n"
repeat_region   27983. .28027
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   complement(28187. .28382)
                /evidence=not_experimental
                /rpt_family="MER20"
repeat_region   28857. .28993
                /evidence=not_experimental
                /rpt_family="FLAM_C"
repeat_region   29511. .29565
                /evidence=not_experimental
                /rpt_family="(TG)n"
repeat_region   29582. .29620
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   29781. .29923
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(30500. .30659)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(30851. .31133)
                /evidence=not_experimental
                /rpt_family="THE1C"
repeat_region   31306. .31348
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region   32634. .32725

  Query Match           2.0%;  Score 27;  DB 9;  Length 105199;
  Best Local Similarity  100.0%;  Pred. No. 0.00053;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db      96086 AGATGTCAAAAAAAAAAAAAAAAAAAA 96112


RESULT 39
AC112138
LOCUS       AC112138              133897 bp    DNA     linear   PRI 29-MAY-2002
DEFINITION  Homo sapiens 3 BAC RP11-715D1 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC112138
VERSION     AC112138.3  GI:21240534
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 133897)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
            Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
            Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
            Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
            Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
            Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
            Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 133897)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 133897)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 133897)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 29, 2002 this sequence version replaced gi:20127825.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

            QUALSTAT-REPORT.
FEATURES                Location/Qualifiers
     source             1. .133897
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="3"
                        /clone="RP11-715D1"
     repeat_region      2057. .2229
                        /rpt_family="MIR"
     repeat_region      2318. .2351
                        /rpt_family="(CA)n"
     repeat_region      2772. .3078
                        /rpt_family="AluY"
     repeat_region      3294. .3585
                        /rpt_family="AluSx"
     repeat_region      5345. .5474
                        /rpt_family="L1MC5"
     repeat_region      6402. .6494
                        /rpt_family="MIR"
     repeat_region      6906. .6929
                        /rpt_family="AT_rich"
     repeat_region      complement(7169. .7337)
                        /rpt_family="AluSq"
     repeat_region      7412. .7457
                        /rpt_family="MADE1"
     repeat_region      7459. .8190
                        /rpt_family="L2"
     repeat_region      8730. .9295
                        /rpt_family="L2"
     repeat_region      9300. .9384
                        /rpt_family="(TA)n"
     repeat_region      9440. .9477
                        /rpt_family="(CA)n"
     repeat_region      10044. .10137
                        /rpt_family="AluY"
     repeat_region      12868. .12894
                        /rpt_family="AT_rich"
     repeat_region      13556. .13770
                        /rpt_family="L2"
     repeat_region      13948. .13981
                        /rpt_family="(TG)n"
     repeat_region      13982. .14012
```

```
                     /rpt_family="(TG)n"
     repeat_region   complement(14141. .14390)
                     /rpt_family="L1ME1"
     repeat_region   14495. .14537
                     /rpt_family="AT_rich"
     repeat_region   complement(15243. .15548)
                     /rpt_family="AluSc"
     repeat_region   complement(16089. .16294)
                     /rpt_family="MER58C"
     repeat_region   complement(16575. .17125)
                     /rpt_family="L1ME1"
     repeat_region   17134. .17286
                     /rpt_family="AluJo/FRAM"
     repeat_region   complement(17289. .17559)
                     /rpt_family="L1ME1"
     repeat_region   complement(17560. .17865)
                     /rpt_family="AluSp"
     repeat_region   complement(17866. .19157)
                     /rpt_family="L1ME1"
     repeat_region   19762. .20121
                     /rpt_family="MLT1A1"
     repeat_region   complement(20939. .21228)
                     /rpt_family="AluSx"
     repeat_region   complement(21907. .22276)
                     /rpt_family="MER46B"
     repeat_region   22633. .22946
                     /rpt_family="AluSx"
     repeat_region   22977. .23247
                     /rpt_family="L2"
     repeat_region   complement(24059. .24161)
                     /rpt_family="MER5A"
     repeat_region   complement(24962. .25164)
                     /rpt_family="U3"
     repeat_region   complement(25225. .25306)
                     /rpt_family="MIR"
     repeat_region   25602. .25817
                     /rpt_family="MIR"
     repeat_region   26057. .26195
                     /rpt_family="MER69A"
     repeat_region   26196. .26219
                     /rpt_family="(CAAAA)n"
     repeat_region   26220. .26233
                     /rpt_family="MER69A"
     repeat_region   26553. .26837
                     /rpt_family="AluY"
     repeat_region   complement(28847. .28982)
                     /rpt_family="L2"

  Query Match             2.0%;  Score 27;  DB 9;  Length 133897;
  Best Local Similarity   100.0%;  Pred. No. 0.00054;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db     123415 AGATGTCAAAAAAAAAAAAAAAAAAAA 123441


RESULT 40
AC099326/c
LOCUS       AC099326               152772 bp    DNA     linear   PRI 24-JAN-2002
DEFINITION  Homo sapiens chromosome 3 clone RP11-7B12, complete sequence.
ACCESSION   AC099326 AC010957
VERSION     AC099326.1  GI:16874860
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152772)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152772)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-2001) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 152772)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Nov 9, 2001 this sequence version replaced gi:10198411.
            -------------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
                Drafting Center: BCM
            -------------- Project Information
                Center project name: chr-3
                Center clone name: RP11-7B12 (bc0603)
            ------------- Summary Statistics
                Sequencing vector: plasmid; 34% of reads
                Sequencing vector: unknown; 66% of reads
                Chemistry: Dye-terminator ET; 57% of reads
                Chemistry: Dye-terminator Big Dye; 43% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 152669 bases at least Q40
                Consensus quality: 152763 bases at least Q30
                Consensus quality: 152772 bases at least Q20
                Insert size: 152772; sum-of-contigs
                Quality coverage: 11.0x in Q20 bases; sum-of-contigs
            --------------.
            Overlapping Sequences:
                5':  RP11-7D10 (UWGC:bc0595) AC010959
                3':  Mapping in progress
            --------------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.
             Base-by-base quality values are not generally visible from the
             GenBank flat file format but are available as part
             of this entry's ASN.1 file.

            This sequence was finished as follows unless otherwise noted:
            all regions were either double-stranded or sequenced with an
            alternate chemistry or covered by high quality data (i.e., Phred
            quality >= 30); an attempt was made to resolve all sequencing
            problems, such as compressions and repeats; all regions were
            covered by at least one plasmid subclone or more than one M13
            subclone; and the assembly was confirmed by restriction digest.
            --------------------
            Sequence Validation:
             This sequence has been validated by Multiple Complete Digest
             fingerprinting. Comparison of the experimentally derived digest
             fragments with sequence-predicted fragments is given below.
             The electronically-digested sequence consists of both insert and
             vector, in order to accurately represent the entire circular BAC.
             Small fragments below a variable cutoff (approximately 400-800 bp)
             are not resolved in the fingerprint and hence do not appear
             in the table. There are no significant remaining discrepancies
             between the experimental and predicted values. Uniquely ordered
             fragments are separated by dashed lines.
                  EcoRI                  HindIII                  BglII
            ------------------   ------------------   ------------------
            SeqDerMap  FngrPrnt   SeqDerMap  FngrPrnt   SeqDerMap  FngrPrnt
            ---------  --------   ---------  --------   ---------  --------

            ---------             ---------             ---------
                 8696       8700       1323       1309       12111      11940
```

6	<800	6382	6519	2067	2160
2532	2547	512	<800	8527	8714
3139	3178	449	<800	1375	1342
4975	5077	1133	1128	893	930
2418	2547	2746	2762	4253	4526
1931	1905	2037	2121	436	<800
7381	7438	2845	2869	925	930
1269	1212	3682	3728	2899	2789
3691	3662	3956	3889	16122	16215
3330	3376	1008	979	2045	2160
8174	8159	5364	5343	3498	3433
5231	5077	2323	2348	650	<800
2843	2906	1357	1309	13741	13713
5157	5077	5192	5120	60	<800
181	<800	10560	10611	5778	6013
1636	1608	758	<800	2421	2399
556	<800	10106	10165	1491	1472
1494	1474	5709	5695	2486	2399
360	<800	342	<800	1861	1904
5006	5077	2678	2762	12854	12717
638	<800	1893	1917	16	<800
5012	5077	304	<800	18	<800
256	<800	3190	3256	1286	1265
6568	6477	2488	2541	3630	3895
414	<800	5865	5695	1037	1057
2045	2079	1457	1442	4025	4293
17911	18009	5628	5695	3747	3895
756	<800	496	<800	6587	6827
5394	5422	886	979	2148	2160
1063	1053	4473	4390	10135	10124
1504	1474	1171	1128	1908	1904
3606	3376	465	<800	719	<800
4911	5077	7137	7196	2175	2160
208	<800	1039	979	2907	2789
5151	5077	234	<800	1064	1057
918	923	4265	4216	3648	3895
4060	4080	934	979	10740	10687
6	<800	6627	6519	2316	2399
3386	3376	1468	1442	3583	3433
675	<800	2168	2121	2782	2789
707	<800	951	979	188	<800
1393	1380	2339	2439	319	<800
8957	8700	2842	2869		
3442	3662	3748	3728		
4672	4584	1164	1128		
2212	2288	3610	3585		
4156	4080	3583	3585		
1444	1474	457	<800		
		7339	7196		
		559	<800		
		1883	1917		
		8831	8854		
		1515	1442		

FEATURES Location/Qualifiers
source 1. .152772
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"

Query Match 2.0%; Score 27; DB 9; Length 152772;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||||
Db 146084 AGATGTCAAAAAAAAAAAAAAAAAAAA 146058

RESULT 41
AC063923/c
LOCUS AC063923 153448 bp DNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens 3 BAC RP11-59E19 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC063923
VERSION AC063923.21 GI:19807681
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 153448)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 153448)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 153448)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 153448)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 153448)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Mar 29, 2002 this sequence version replaced gi:19718593.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1. .153448
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-59E19"
misc_feature	1. .2019
	/note="overlaps bases 73308. .75323 of clone AC076971"
	/function="clone overlap"
repeat_region	complement(133. .416)
	/rpt_family="AluJb"
repeat_region	419. .471
	/rpt_family="L1MC5"
repeat_region	complement(472. .633)
	/rpt_family="FRAM"
repeat_region	634. .718
	/rpt_family="L1MC5"
repeat_region	780. .1157
	/rpt_family="MLT1D"
repeat_region	complement(1434. .2333)
	/rpt_family="L1M4"
repeat_region	complement(2402. .2577)
	/rpt_family="L1M4"
repeat_region	2583. .2624
	/rpt_family="(TA)n"
repeat_region	complement(2625. .2872)
	/rpt_family="AluSx"
repeat_region	2873. .2900
	/rpt_family="(TA)n"
repeat_region	complement(2961. .3085)
	/rpt_family="L1M4"
repeat_region	complement(3182. .3481)
	/rpt_family="AluJo"
repeat_region	complement(4166. .4456)
	/rpt_family="AluSx"
repeat_region	4686. .4865
	/rpt_family="(TTCTCC)n"
repeat_region	complement(5577. .5875)
	/rpt_family="AluY"
repeat_region	complement(5879. .6244)
	/rpt_family="MLT1E"
repeat_region	complement(6339. .6432)
	/rpt_family="MLT1E"
repeat_region	6581. .7060
	/rpt_family="LTR40a"
repeat_region	7537. .7566
	/rpt_family="(TTTTA)n"
repeat_region	complement(7567. .7714)
	/rpt_family="AluSx"

```
     repeat_region   complement(7715. .8013)
                     /rpt_family="AluSc"
     repeat_region   complement(8014. .8155)
                     /rpt_family="AluSx"
     repeat_region   complement(8771. .8835)
                     /rpt_family="MIR"
     repeat_region   complement(8836. .9096)
                     /rpt_family="AluJb"
     repeat_region   complement(9097. .9239)
                     /rpt_family="MIR"
     repeat_region   9443. .9889
                     /rpt_family="L2"
     repeat_region   9911. .9989
                     /rpt_family="L2"
     repeat_region   9990. .10032
                     /rpt_family="(TGAA)n"
     repeat_region   complement(10061. .10148)
                     /rpt_family="L2"
     repeat_region   10149. .10297
                     /rpt_family="AluJo"
     repeat_region   10298. .10606
                     /rpt_family="AluSc"
     repeat_region   10607. .10765
                     /rpt_family="AluJo"
     repeat_region   10786. .11020
                     /rpt_family="HAL1"
     repeat_region   11125. .11448
                     /rpt_family="HAL1"
     repeat_region   12498. .12975
                     /rpt_family="MLT2B"
     repeat_region   complement(12976. .13186)
                     /rpt_family="AluSg/x"
     repeat_region   13187. .13263

  Query Match             2.0%;  Score 27;  DB 9;  Length 153448;
  Best Local Similarity   100.0%;  Pred. No. 0.00055;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db     152091 AGATGTCAAAAAAAAAAAAAAAAAAAA 152065


RESULT 42
AC104449/c
LOCUS       AC104449               174531 bp    DNA     linear   PRI 10-JUN-2003
DEFINITION  Homo sapiens chromosome 3 clone RP11-460N16, complete sequence.
ACCESSION   AC104449
VERSION     AC104449.3  GI:31560233
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 174531)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and
            Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174531)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-2001) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 174531)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
            and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   4  (bases 1 to 174531)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and
            Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-2003) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Jun 10, 2003 this sequence version replaced gi:22004189.
            ------------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
            ------------- Project Information
                Center project name: chr-3
                Center clone name: RP11-460N16 (bc0764)
            ------------ Summary Statistics
                Sequencing vector: plasmid; 100% of reads
                Chemistry: Dye-terminator ET; 89% of reads
                Chemistry: Dye-terminator Big Dye; 11% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 173867 bases at least Q40
                Consensus quality: 174438 bases at least Q30
                Consensus quality: 174513 bases at least Q20
                Insert size: 174531; sum-of-contigs
                Quality coverage: 9.9x in Q20 bases; sum-of-contigs
            -------------.
            Overlapping Sequences:
                5': RP11-7B12 (UWGC:bc0603) AC099326, 61951-bp overlap
                3': RP11-215K24 (UWGC:bc0633) AC099666, 86241-bp overlap
            --------------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.
             Base-by-base quality values are not generally visible from the
             GenBank flat file format but are available as part
             of this entry's ASN.1 file.

            This sequence was finished as follows unless otherwise noted:
            all regions were either double-stranded or sequenced with an
            alternate chemistry or covered by high quality data (i.e., Phred
            quality >= 30); an attempt was made to resolve all sequencing
            problems, such as compressions and repeats; all regions were
            covered by at least one plasmid subclone or more than one M13
            subclone; and the assembly was confirmed by restriction digest.
            --------------------
            Sequence Validation:
             This sequence has been validated by Multiple Complete Digest
             fingerprinting. Comparison of the experimentally derived digest
             fragments with sequence-predicted fragments is given below.
             The electronically-digested sequence consists of both insert and
             vector, in order to accurately represent the entire circular BAC.
             Small fragments below a variable cutoff (approximately 400-800 bp)
             are not resolved in the fingerprint and hence do not appear
             in the table. There are no significant remaining discrepancies
             between the experimental and predicted values. Uniquely ordered
             fragments are separated by dashed lines.
```

BglII		EcoRI		HindIII	
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
3466	3516	8696	8744	1508	1475
2067	2067	6	<800	6382	6533
8045	8044	1149	1185	512	<800
462	<800	3436	3418	449	<800

3573	3730	314	<800	1586	1577
3957	4016	611	<800	6006	6016
231	<800	1206	1185	2640	2612
5379	5264	11764	11362	989	1000
208	<800	645	<800	2203	2351
541	<800	2775	2642	6314	6297
3765	3910	533	<800	5814	5690
4033	4185	4992	4944	773	<800
604	<800	3290	3297	2358	2351
541	<800	2896	2958	3526	3509
6856	6776	6623	6599	1475	1475
9840	9668	1188	1185	5668	5690
479	<800	2656	2642	4882	4809
1716	1678	529	<800	2728	2806
6109	6092	10269	10030	998	1000
913	919	186	<800	2264	2351
4352	4330	5513	5517	7957	7979
1536	1530	2479	2486	3868	3860
2338	2345	8122	8148	2347	2351
16968	17243	5075	5133	2766	2806
9	<800	2032	2025	1778	1790
624	<800	747	765	1202	1180
5148	5070	845	868	7646	7689
564	<800	6360	6444	6107	6016
2737	2743	486	<800	2081	2076
6213	6092	16691	16973	61	<800
1102	1115	1625	1613	223	<800
8039	8044	4695	4944	4130	4049
688	<800	2848	2844	274	<800
4877	4841	2532	2642	306	<800
1671	1678	3140	3156	4169	4049
1685	1678	4975	4944	3560	3655
976	919	2418	2486	9213	9229
4328	4330	1931	1908	801	822
1376	1335	7381	7416	2192	2351
893	919	1269	1185	1617	1577
4253	4330	3691	3700	1191	1180
436	<800	3332	3297	2746	2806
925	919	8174	8148	2038	2076
2899	2914	5231	5133	2845	2806
16124	16189	2843	2844	3682	3860
2045	2067	5157	5133	3956	4049
3498	3516	181	<800	1008	1000
650	<800	1636	1613	5364	5318
13741	13451	556	<800	2323	2351
60	<800	1494	1474	1357	1329
5778	5725	360	<800	5194	5136
2421	2458	5006	5133	10560	10559
1491	1468	641	<800	758	795
				10106	10067

Query Match 2.0%; Score 27; DB 9; Length 174531;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db      55262 AGATGTCAAAAAAAAAAAAAAAAAAAA 55236
```

RESULT 43
AC010959
LOCUS AC010959 177672 bp DNA linear HTG 08-MAR-2000
DEFINITION Homo sapiens clone RP11-7D10, WORKING DRAFT SEQUENCE, 25 unordered pieces.
ACCESSION AC010959
VERSION AC010959.3 GI:7209492
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 177672)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-7D10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 177672)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 8, 2000 this sequence version replaced gi:6978150.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
    Center: Whitehead Institute/ MIT Center for Genome Research
    Center code: WIBR
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
    Center project name: L2922
    Center clone name: 7_D_10
-------------- Summary Statistics
    Sequencing vector: M13; M77815; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 149787 bases at least Q40
    Consensus quality: 162330 bases at least Q30
    Consensus quality: 169351 bases at least Q20
    Insert size: 182000; agarose-fp
    Insert size: 175272; sum-of-contigs
    Quality coverage: 3.9 in Q20 bases; agarose-fp
    Quality coverage: 4.1 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*         1     1129: contig of 1129 bp in length
*      1130     1229: gap of 100 bp
*      1230     2365: contig of 1136 bp in length
*      2366     2465: gap of 100 bp
*      2466     4110: contig of 1645 bp in length
*      4111     4210: gap of 100 bp
*      4211     5970: contig of 1760 bp in length
*      5971     6070: gap of 100 bp
*      6071     7611: contig of 1541 bp in length
*      7612     7711: gap of 100 bp
*      7712     9717: contig of 2006 bp in length
*      9718     9817: gap of 100 bp
*      9818    12688: contig of 2871 bp in length
*     12689    12788: gap of 100 bp
*     12789    16615: contig of 3827 bp in length
*     16616    16715: gap of 100 bp
*     16716    20067: contig of 3352 bp in length
*     20068    20167: gap of 100 bp
*     20168    23688: contig of 3521 bp in length
*     23689    23788: gap of 100 bp
*     23789    27981: contig of 4193 bp in length
*     27982    28081: gap of 100 bp
*     28082    31601: contig of 3520 bp in length
*     31602    31701: gap of 100 bp
*     31702    37510: contig of 5809 bp in length
*     37511    37610: gap of 100 bp
*     37611    42701: contig of 5091 bp in length
*     42702    42801: gap of 100 bp
*     42802    50137: contig of 7336 bp in length
*     50138    50237: gap of 100 bp
*     50238    57522: contig of 7285 bp in length
*     57523    57622: gap of 100 bp
*     57623    66946: contig of 9324 bp in length
*     66947    67046: gap of 100 bp
*     67047    74537: contig of 7491 bp in length
*     74538    74637: gap of 100 bp
*     74638    78928: contig of 4291 bp in length
*     78929    79028: gap of 100 bp
*     79029    89116: contig of 10088 bp in length
*     89117    89216: gap of 100 bp
*     89217    98944: contig of 9728 bp in length
*     98945    99044: gap of 100 bp
*     99045   113038: contig of 13994 bp in length
*    113039   113138: gap of 100 bp
*    113139   131560: contig of 18422 bp in length
*    131561   131660: gap of 100 bp
*    131661   150982: contig of 19322 bp in length
*    150983   151082: gap of 100 bp
*    151083   177672: contig of 26590 bp in length.
FEATURES             Location/Qualifiers
     source          1. .177672
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-7D10"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1129
                     /note="assembly_fragment"
     misc_feature    1230. .2365
                     /note="assembly_fragment"
     misc_feature    2466. .4110
                     /note="assembly_fragment"
     misc_feature    4211. .5970
                     /note="assembly_fragment"
     misc_feature    6071. .7611
                     /note="assembly_fragment"
     misc_feature    7712. .9717
                     /note="assembly_fragment"
     misc_feature    9818. .12688
                     /note="assembly_fragment"
     misc_feature    12789. .16615
                     /note="assembly_fragment"
     misc_feature    16716. .20067
                     /note="assembly_fragment"
     misc_feature    20168. .23688
                     /note="assembly_fragment"
     misc_feature    23789. .27981
                     /note="assembly_fragment"
     misc_feature    28082. .31601
                     /note="assembly_fragment"
     misc_feature    31702. .37510
                     /note="assembly_fragment"
     misc_feature    37611. .42701
                     /note="assembly_fragment"
     misc_feature    42802. .50137
                     /note="assembly_fragment"
     misc_feature    50238. .57522
                     /note="assembly_fragment"
     misc_feature    57623. .66946
                     /note="assembly_fragment"
     misc_feature    67047. .74537
                     /note="assembly_fragment"
     misc_feature    74638. .78928
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    79029. .89116
                     /note="assembly_fragment"
     misc_feature    89217. .98944
                     /note="assembly_fragment"
     misc_feature    99045. .113038
                     /note="assembly_fragment"
     misc_feature    113139. .131560
                     /note="assembly_fragment"
     misc_feature    131661. .150982
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    151083. .177672
                     /note="assembly_fragment"
ORIGIN

  Query Match             2.0%;  Score 27;  DB 2;  Length 177672;
  Best Local Similarity   100.0%;  Pred. No. 0.00055;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
```

```
                    ||||||||||||||||||||||||||
Db     102188 AGATGTCAAAAAAAAAAAAAAAAAAAA 102214


RESULT 44
AC106783
LOCUS       AC106783               179585 bp    DNA     linear   PRI 14-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-347K10, complete sequence.
ACCESSION   AC106783
VERSION     AC106783.3  GI:28951177
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 179585)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 179585)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 179585)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 179585)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 14, 2003 this sequence version replaced gi:18642717.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.7.
FEATURES             Location/Qualifiers
     source          1. .179585
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-347K10"
ORIGIN

  Query Match             2.0%;  Score 27;  DB 9;  Length 179585;
  Best Local Similarity   100.0%;  Pred. No. 0.00055;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db      13413 AGATGTCAAAAAAAAAAAAAAAAAAAA 13439


RESULT 45
AX706971
LOCUS       AX706971               189013 bp    DNA     linear   PAT 04-APR-2003
DEFINITION  Sequence 669 from Patent WO03013534.
ACCESSION   AX706971
VERSION     AX706971.1  GI:29563296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Heinrich,G. and Kerb,R.
  TITLE     Methods for the treatment of cancer with irinotecan based on CYP3A5
  JOURNAL   Patent: WO 03013534-A 669 20-FEB-2003;
            Epidauros Biotechnologie AG (DE)
FEATURES             Location/Qualifiers
     source          1. .189013
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             2.0%;  Score 27;  DB 6;  Length 189013;
  Best Local Similarity   100.0%;  Pred. No. 0.00056;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db     123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963


RESULT 46
AX707901
LOCUS       AX707901               189013 bp    DNA     linear   PAT 04-APR-2003
DEFINITION  Sequence 669 from Patent WO03013536.
ACCESSION   AX707901
VERSION     AX707901.1  GI:29563976
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Heinrich,G. and Kerb,R.
  TITLE     Methods for treatment of cancer using irinotecan based on UGT1A1
  JOURNAL   Patent: WO 03013536-A 669 20-FEB-2003;
            Epidauros Biotechnologie AG (DE)
FEATURES             Location/Qualifiers
     source          1. .189013
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             2.0%;  Score 27;  DB 6;  Length 189013;
  Best Local Similarity   100.0%;  Pred. No. 0.00056;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db     123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963


RESULT 47
AC025277
LOCUS       AC025277               189032 bp    DNA     linear   PRI 25-NOV-2003
DEFINITION  Homo sapiens chromosome 16 clone CTD-3229J4, complete sequence.
ACCESSION   AC025277
VERSION     AC025277.10  GI:38502330
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 189032)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 189032)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAR-2000) Production Sequencing Facility, DOE Joint
```

```
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 189032)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 189032)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 189032)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-SEP-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   6  (bases 1 to 189032)
  AUTHORS   Stanford Human Genome Center and Los Alamos National Laboratory.
  CONSRTM   DOE Joint Genome Institute
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Nov 25, 2003 this sequence version replaced gi:22725948.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .189032
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-3229J4"
ORIGIN

  Query Match             2.0%;  Score 27;  DB 9;  Length 189032;
  Best Local Similarity   100.0%;  Pred. No. 0.00056;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||
Db     123946 AGATGTCAAAAAAAAAAAAAAAAAAAA 123972


RESULT 48
AC110206/c
LOCUS       AC110206               199045 bp    DNA     linear   HTG 17-DEC-2003
DEFINITION  Mus musculus chromosome 19 clone RP23-155H17 map 19, *** SEQUENCING
            IN PROGRESS ***, 2 ordered pieces.
ACCESSION   AC110206
VERSION     AC110206.18  GI:39841427
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 199045)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 19, clone RP23-155H17
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199045)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 199045)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Dec 15, 2003 this sequence version replaced gi:37202218.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L20741
                Center clone name: 155_H_17
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   136971: contig of 136971 bp in length
            *  136972   137071: gap of 100 bp
            *  137072   199045: contig of 61974 bp in length.
FEATURES             Location/Qualifiers
     source          1. .199045
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="19"
                     /map="19"
```

```
                     /clone="RP23-155H17"
                     /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

  Query Match              2.0%;  Score 27;  DB 2;  Length 199045;
  Best Local Similarity   100.0%;  Pred. No. 0.00056;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db      68595 AGATGTCAAAAAAAAAAAAAAAAAAAA 68569


RESULT 49
AC133100/c
LOCUS       AC133100               204855 bp    DNA     linear   ROD 25-NOV-2003
DEFINITION  Mus musculus BAC clone RP23-98O11 from chromosome 19, complete
            sequence.
ACCESSION   AC133100
VERSION     AC133100.5  GI:37806568
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 204855)
  AUTHORS   Harkins,R., Haglund,K. and Haakenson,W.
  TITLE     The sequence of Mus musculus BAC clone RP23-98O11
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 204855)
  AUTHORS   Wilson,R.
  TITLE     Sequencing of Mus musculus
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 204855)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 204855)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 204855)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   6  (bases 1 to 204855)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 22, 2003 this sequence version replaced gi:33457295.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BA0098O11
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
            and Minako Tateno in the laboratory of Pieter de Jong
            (http://www.chori.org) from female C57BL/6J mouse kidney and/or
            brain genomic DNA.  The clone and detailed information can be
            obtained from Research Genetics, Inc. (http://www.resgen.com) or
            Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone. This clone is
            overlapped by AC130544.
FEATURES             Location/Qualifiers
     source          1. .204855
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="19"
                     /map="19"
                     /clone="RP23-98O11"
                     /clone_lib="RPCI-23"
     repeat_region   5. .65
                     /rpt_family="B4"
     repeat_region   1412. .1604
                     /rpt_family="B2"
     repeat_region   2198. .2326
                     /rpt_family="L1"
     repeat_region   2320. .2597
                     /rpt_family="L1"
     repeat_region   2644. .2722
                     /rpt_family="B4"
     repeat_region   2723. .2842
                     /rpt_family="L1"
     repeat_region   2847. .3068
                     /rpt_family="L1"
     repeat_region   3115. .3323
                     /rpt_family="B2"
     repeat_region   3432. .4159
                     /rpt_family="L1"
     repeat_region   5707. .5863
                     /rpt_family="ERVL"
     repeat_region   6013. .6293
                     /rpt_family="B4"
     repeat_region   6947. .7137
                     /rpt_family="MER1_type"
     repeat_region   7516. .7715
                     /rpt_family="B2"
     repeat_region   7978. .8084
                     /rpt_family="L1"
     repeat_region   8623. .8766
                     /rpt_family="ERVK"
     repeat_region   9230. .9365
                     /rpt_family="MIR"
     repeat_region   11689. .11827
                     /rpt_family="MIR"
     misc_feature    12154. .13463
                     /note="CpG island (%GC=62.3, o/e=0.91, #CpGs=117)"
     repeat_region   14778. .14874
                     /rpt_family="Alu"
     repeat_region   15739. .16103
                     /rpt_family="MaLR"
     repeat_region   17882. .18040
                     /rpt_family="MER1_type"
     repeat_region   18287. .18371
                     /rpt_family="ID"
```

```
repeat_region   18556. .18866
                /rpt_family="B4"
repeat_region   20123. .20175
                /rpt_family="5S"
repeat_region   20397. .20697
                /rpt_family="MER1_type"
repeat_region   20899. .21004
                /rpt_family="MER1_type"
repeat_region   21093. .21197
                /rpt_family="MER2_type"
repeat_region   22734. .23067
                /rpt_family="RMER12"
repeat_region   23155. .23221
                /rpt_family="RMER12"
repeat_region   25357. .25466
                /rpt_family="Alu"
repeat_region   25637. .25787
                /rpt_family="MER1_type"
repeat_region   28526. .28927
                /rpt_family="L1"
repeat_region   29467. .29563
                /rpt_family="Alu"
repeat_region   31412. .31561
                /rpt_family="Alu"
repeat_region   31737. .31839
                /rpt_family="Alu"
repeat_region   31993. .32138
                /rpt_family="Alu"
repeat_region   32744. .32943
                /rpt_family="MER1_type"
repeat_region   33577. .33663
                /rpt_family="Alu"
repeat_region   33610. .33750
                /rpt_family="B4"
repeat_region   36229. .36350
                /rpt_family="Alu"
repeat_region   36679. .36762
                /rpt_family="MIR"
repeat_region   37830. .37953
                /rpt_family="MER1_type"
repeat_region   38461. .38665
                /rpt_family="B2"
repeat_region   38950. .39327
                /rpt_family="MaLR"
repeat_region   39497. .39639
                /rpt_family="B4"
repeat_region   40408. .40684
                /rpt_family="B4"
repeat_region   41391. .42306
                /rpt_family="ERVK"
repeat_region   43473. .43832
                /rpt_family="L1"
repeat_region   44689. .44835
                /rpt_family="Tc2"
repeat_region   45866. .46200
                /rpt_family="MaLR"
repeat_region   46640. .46786
                /rpt_family="Alu"
repeat_region   51588. .51858
                /rpt_family="MaLR"
repeat_region   52268. .52352
                /rpt_family="ID"
tRNA            52276. .52347
                /product="tRNA-Ala"
                /note="Likely pseudogene (HMM Sc=16.12 / Sec struct
                Sc=4.29)"
repeat_region   54145. .54483
                /rpt_family="MaLR"
repeat_region   54657. .54860

Query Match             2.0%;  Score 27;  DB 10;  Length 204855;
Best Local Similarity   100.0%;  Pred. No. 0.00056;
Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db     197260 AGATGTCAAAAAAAAAAAAAAAAAAAA 197234


RESULT 50
AC099108
LOCUS       AC099108               232802 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-186G8, WORKING DRAFT SEQUENCE.
ACCESSION   AC099108
VERSION     AC099108.7  GI:30522525
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 232802)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 232802)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 232802)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:23096531.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GGPQ
                Center clone name: CH230-186G8
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 224396 bases at least Q40
                Consensus quality: 225844 bases at least Q30
                Consensus quality: 226943 bases at least Q20
                Estimated insert size: 233584; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   232802: contig of 232802 bp in length.
FEATURES            Location/Qualifiers
     source         1. .232802
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-186G8"
     misc_feature   1. .1210
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   1261. .3361
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   4020. .6873
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   8435. .8986
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH292506"
     misc_feature   complement(230253. .230668)
                    /note="clone_boundary
                    clone_end:T7
                    site:EcoRI
                    end_sequence:BH292505"
ORIGIN

  Query Match             2.0%;  Score 27;  DB 2;  Length 232802;
  Best Local Similarity   100.0%;  Pred. No. 0.00057;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db      74169 AGATGTCAAAAAAAAAAAAAAAAAAAA 74195


Search completed: September 17, 2004, 04:46:08
Job time : 5170 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 20:16:53 ; Search time 552 Seconds
(without alignments)
10212.607 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtggtcacaggtag.....aaaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1327	100.0	3826	6	ABK12806	Abk12806 Human cDN
	2	1308	98.6	5858	4	AAK80624	Aak80624 Human imm
	3	1308	98.6	22893	6	ABK12810	Abk12810 Human tum
	4	1307	98.5	3243	9	ADC37268	Adc37268 Nuclear f
	5	1257	94.7	5866	4	AAK80625	Aak80625 Human imm
	6	1257	94.7	30676	6	ABK12811	Abk12811 Human tum
	7	1256	94.6	3436	4	AAH14509	Aah14509 Human cDN
	8	726	54.7	5858	4	AAK80626	Aak80626 Human imm
	9	726	54.7	30625	6	ABK12808	Abk12808 Human tum
	10	452	34.1	2207	9	ADC37454	Adc37454 Nuclear f
	11	452	34.1	2207	9	ADC37264	Adc37264 Nuclear f
	12	452	34.1	2246	9	ADC37266	Adc37266 Nuclear f
c	13	423	31.9	557	4	AAH11818	Aah11818 Human cDN
	14	60	4.5	60	6	ABN43449	Abn43449 Human spl
	15	29	2.2	614	6	ABN63808	Abn63808 Human can
	16	27	2.0	470	3	AAC59529	Aac59529 Human sec
	17	27	2.0	189013	7	ACF62741	Acf62741 Cancer ba
	18	27	2.0	189013	7	ADB20856	Adb20856 MRP1 base
	19	27	2.0	189013	9	ADB87945	Adb87945 Human UGT
	20	27	2.0	189013	9	ADB96928	Adb96928 Human MDR
	21	27	2.0	189013	9	ADB92119	Adb92119 Human MDR
	22	26	2.0	273	6	ABQ60500	Abq60500 Human col
c	23	26	2.0	393	7	ABX41090	Abx41090 Bovine ES
c	24	26	2.0	411	7	ABX48150	Abx48150 Bovine ES
c	25	26	2.0	659	4	AAL24218	Aal24218 Human bre
	26	26	2.0	729	2	AAQ35208	Aaq35208 Sequence
	27	26	2.0	2295	4	AAS02396	Aas02396 Human sec
	28	26	2.0	4702	7	ABX76356	Abx76356 Lung canc
	29	26	2.0	110000	3	AAF22305_07	Continuation (8 of
	30	25	1.9	184	6	ABL86275	Abl86275 Human ova
c	31	25	1.9	191	6	ABL65710	Abl65710 Lung canc
c	32	25	1.9	191	6	ABL64256	Abl64256 Stomach c
c	33	25	1.9	191	6	ABN94027	Abn94027 Gene #525
	34	25	1.9	207	6	ABL87801	Abl87801 Human ova
c	35	25	1.9	222	6	ABN94089	Abn94089 Gene #587
c	36	25	1.9	330	4	AAS56440	Aas56440 Human cDN
c	37	25	1.9	354	7	ABX36291	Abx36291 Bovine ES
c	38	25	1.9	365	4	AAL09871	Aal09871 Human bre
	39	25	1.9	390	7	ACC48717	Acc48717 Mouse wou
c	40	25	1.9	414	4	AAL18543	Aal18543 Human bre
	41	25	1.9	425	5	ABV18188	Abv18188 Human pro
	42	25	1.9	430	8	ACH20487	Ach20487 Human adu
	43	25	1.9	433	7	ACF03901	Acf03901 Rice endo
c	44	25	1.9	464	7	ABX43422	Abx43422 Bovine ES
	45	25	1.9	469	2	AAQ74392	Aaq74392 Clone pSP
	46	25	1.9	469	2	AAQ72957	Aaq72957 Leech PCR
	47	25	1.9	469	2	AAV33944	Aav33944 Leech FXa
	48	25	1.9	469	2	AAX01821	Aax01821 H. medici
	49	25	1.9	469	2	AAV55665	Aav55665 Factor Xa
	50	25	1.9	469	4	AAD04138	Aad04138 DNA clone
	51	25	1.9	527	3	AAC78393	Aac78393 Human can
	52	25	1.9	549	6	ABQ60567	Abq60567 Human col
c	53	25	1.9	558	3	AAC77467	Aac77467 Human ORF
	54	25	1.9	591	4	AAS41534	Aas41534 cDNA enco
	55	25	1.9	591	4	AAL01169	Aal01169 Human rep
	56	25	1.9	591	4	ABL96630	Abl96630 Human tes
	57	25	1.9	609	5	ABV54411	Abv54411 Human pro
	58	25	1.9	645	6	ABK34272	Abk34272 Human cDN
c	59	25	1.9	657	6	ABT08966	Abt08966 Phase-1 R
	60	25	1.9	677	4	AAS40979	Aas40979 cDNA enco
	61	25	1.9	716	2	AAZ17556	Aaz17556 Human gen
c	62	25	1.9	748	4	AAI95890	Aai95890 Human neu
	63	25	1.9	867	2	AAT66791	Aat66791 Partial s
	64	25	1.9	877	5	AAS76330	Aas76330 DNA encod
	65	25	1.9	1092	3	AAC77923	Aac77923 Human can
	66	25	1.9	1104	9	ADE07313	Ade07313 Novel cod
	67	25	1.9	1451	4	AAH33335	Aah33335 Human col
	68	25	1.9	1452	3	AAC98173	Aac98173 Human col
	69	25	1.9	1539	9	ADC49304	Adc49304 DNA encod
	70	25	1.9	1624	4	AAS25947	Aas25947 Human cDN
	71	25	1.9	1624	7	ABX73288	Abx73288 Human nov
	72	25	1.9	1670	5	ABV26918	Abv26918 Human pro
	73	25	1.9	1670	5	ABV21075	Abv21075 Human pro
	74	25	1.9	1696	3	AAF16214	Aaf16214 Human pro
c	75	25	1.9	1790	6	ABQ98953	Abq98953 Human ORF
	76	25	1.9	1824	4	AAS41354	Aas41354 cDNA enco
	77	25	1.9	1868	8	AAL62037	Aal62037 Human cel
	78	25	1.9	2104	4	AAC87120	Aac87120 Nucleotid
	79	25	1.9	2522	2	AAV82781	Aav82781 Clone df3
	80	25	1.9	2522	6	ABQ92018	Abq92018 Human pol
c	81	25	1.9	2586	9	ADB62721	Adb62721 Human cDN
	82	25	1.9	2629	4	AAH34824	Aah34824 Human col
c	83	25	1.9	2647	5	ABV23296	Abv23296 Human pro
c	84	25	1.9	2647	5	ABV25449	Abv25449 Human pro
c	85	25	1.9	2647	5	ABV30164	Abv30164 Human pro
c	86	25	1.9	2647	5	ABV25571	Abv25571 Human pro
c	87	25	1.9	2647	5	ABV29143	Abv29143 Human pro
	88	25	1.9	2796	2	AAX35727	Aax35727 cDNA enco
	89	25	1.9	3858	9	ADD69641	Add69641 Human REM
	90	25	1.9	3891	3	AAC76424	Aac76424 Human ORF
	91	25	1.9	4493	8	ACA98971	Aca98971 cDNA enco
	92	25	1.9	4498	8	ADA38053	Ada38053 DNA relat
	93	25	1.9	4572	6	ABQ88201	Abq88201 Human ost
	94	25	1.9	4585	4	ABL15850	Abl15850 Drosophil
	95	25	1.9	4717	6	ABK99971	Abk99971 DNA encod
	96	25	1.9	4729	2	AAQ04043	Aaq04043 Recombina

97	25	1.9	4740	3	AAA35032	Aaa35032 Human ade
98	25	1.9	4740	3	AAF21154	Aaf21154 Human low
99	25	1.9	4740	6	ABK50046	Abk50046 cDNA enco
100	25	1.9	4740	6	ABK83766	Abk83766 Human cDN
101	25	1.9	4740	7	ABZ96848	Abz96848 Human nuc
102	25	1.9	4740	7	ABV76142	Abv76142 Human Mac
103	25	1.9	4740	9	ADD25614	Add25614 Binding d
104	25	1.9	5307	8	ADA38055	Ada38055 DNA relat
105	25	1.9	5363	8	ADA38051	Ada38051 DNA relat
106	25	1.9	5371	7	ABZ35979	Abz35979 Human sec
107	25	1.9	5684	9	ADE72685	Ade72685 Human end
108	25	1.9	5808	7	ABX76355	Abx76355 Lung canc
c 109	25	1.9	6405	5	AAF97850	Aaf97850 Human neu
110	25	1.9	7047	3	AAA35034	Aaa35034 Human ade
111	25	1.9	7047	3	AAF21156	Aaf21156 Human low
112	25	1.9	7047	7	ABZ96850	Abz96850 Human nuc
c 113	25	1.9	15765	6	AAD32327	Aad32327 Human NFK
c 114	25	1.9	15765	6	AAD32179	Aad32179 Human NFK
115	25	1.9	18099	8	AAD58283	Aad58283 Human tum
116	25	1.9	27884	4	AAK77781	Aak77781 Human imm
c 117	25	1.9	79528	6	AAL50814	Aal50814 Human can
c 118	25	1.9	110000	6	ABS55320_0	Abs55320 Human tra
c 119	25	1.9	110000	6	ABS55320_3	Continuation (4 of
120	25	1.9	226475	8	AAD58279	Aad58279 Human tum
c 121	25	1.9	341511	6	ABS55200	Abs55200 Genomic D
c 122	24	1.8	40	4	AAH20345	Aah20345 HHV6 viru
c 123	24	1.8	42	4	AAC88926	Aac88926 Mycobacte
c 124	24	1.8	90	2	AAQ98109	Aaq98109 Comb body
c 125	24	1.8	92	6	ABV96450	Abv96450 Human pan
126	24	1.8	158	6	ABL86941	Abl86941 Human ova
127	24	1.8	166	4	AAH71409	Aah71409 Human cer
128	24	1.8	173	7	ABX46137	Abx46137 Bovine ES
129	24	1.8	203	5	ABV19399	Abv19399 Human pro
c 130	24	1.8	206	6	ABL69722	Abl69722 Prostate
c 131	24	1.8	206	6	ABN96234	Abn96234 Gene #273
c 132	24	1.8	229	7	ABX54617	Abx54617 Bovine ES
133	24	1.8	274	4	AAF59571	Aaf59571 Rat urote
c 134	24	1.8	282	5	ABV14508	Abv14508 Human pro
135	24	1.8	289	4	AAS23972	Aas23972 Human ova
136	24	1.8	289	5	AAH82537	Aah82537 Human ova
c 137	24	1.8	299	2	AAV69628	Aav69628 Human sec
138	24	1.8	312	5	ABA14476	Aba14476 Human ner
139	24	1.8	314	5	ABV06723	Abv06723 Human pro
140	24	1.8	333	5	ABV49172	Abv49172 Human pro
c 141	24	1.8	338	4	AAH09999	Aah09999 Human cDN
c 142	24	1.8	361	7	ABZ88882	Abz88882 Human nuc
143	24	1.8	368	6	ABT08933	Abt08933 Phase-1 R
c 144	24	1.8	372	6	ABL84265	Abl84265 Human ova
c 145	24	1.8	373	6	ABQ85428	Abq85428 Arabidops
146	24	1.8	379	5	ABV18002	Abv18002 Human pro
147	24	1.8	386	4	AAI92258	Aai92258 Human pol
148	24	1.8	390	5	ABV54455	Abv54455 Human pro
149	24	1.8	391	5	ABV47791	Abv47791 Human pro
150	24	1.8	393	5	ABV36670	Abv36670 Human pro

ALIGNMENTS

RESULT 1
ABK12806
ID ABK12806 standard; cDNA; 3826 BP.
XX
AC ABK12806;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cDNA encoding tumour suppressor CAR-1.
XX
KW Human; ss; tumour suppressor; gene; CAR-1; cytostatic; cancer; tumour;
KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
KW colon cancer; stomach cancer; breast cancer; endometrial cancer;
KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;
KW head and neck cancer; oesophageal cancer; bone marrow cancer;
KW chromosome 1p31-1p36.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 604. .2031
FT /*tag= a
FT /product= "CAR-1"
XX
PN WO200212285-A2.
XX
PD 14-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-US025269.
XX
PR 10-AUG-2000; 2000US-0225033P.
PR 23-AUG-2000; 2000US-0227560P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Killary A, Chandler D, Lott S;
XX
DR WPI; 2002-269088/31.
DR P-PSDB; AAU78657.
XX
PT New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
PT diagnosing cancer, for altering the phenotype of a tumor cell, for
PT treating cancers or as a diagnostic or prognostic indicator of cancer.
XX
PS Claim 2; Page 134-135; 185pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide being tumour suppressor, CAR-1. Also included are fragments
CC of the polynucleotide from 15-5000 nucleotides, fragments of the protein
CC from 10-50 amino acids, an expression cassette comprising the
CC polynucleotide under the control of a promoter operable in eukaryotic
CC cells, a method for suppressing growth of a cancer cells by contacting
CC the cells with the expression cassette (i.e. gene therapy), a cell
CC comprising the expression cassette, an anti-CAR-1 monoclonal or
CC polyclonal antibody, a hybridoma cell that produces the monoclonal
CC antibody, a method of diagnosing a cancer by assessing the expression of
CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,
CC methods for altering the phenotype of a tumour cell, methods for treating
CC a subject with cancer by administering the tumour suppressor CAR-1, or by
CC administering a nucleic acid encoding the tumour suppressor CAR-1 and a
CC promoter active in eukaryotic cells, where the promoter is operably
CC linked to the region encoding the tumour suppressor, a non-human
CC transgenic eukaryote lacking a functional CAR-1 gene, a non-human
CC transgenic eukaryote that over-expresses CAR-1 as compared to a similar
CC non-transgenic eukaryote, a method of screening a candidate substance for
CC anti-tumour activity by contacting a cell lacking functional CAR-1
CC polypeptide, with a candidate substance and determining the effect of the
CC candidate substance on the cell, an anti-tumour composition produced by
CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate
CC substance, determining the effect of the candidate substance on the cell,
CC identifying a candidate inhibitor substance, and making a composition and
CC an isolated and purified nucleic acid that hybridizes, under high
CC stringency conditions, to a DNA segment comprising about 15-3826 bases of
CC the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
CC altering the phenotype of a tumour cell, for treating cancers (e.g.
CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,
CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
CC other tissues), and as a diagnostic or prognostic indicator of cancer.
CC CAR-1 may also be used in screening compounds for activity in either
CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
CC effect of mutant CAR-1 molecule. The gene for CAR-1 is located on
CC chromosome 1 (1p31-1p36). The present sequence is the cDNA encoding CAR-1
XX
SQ Sequence 3826 BP; 797 A; 1286 C; 999 G; 744 T; 0 U; 0 Other;

Query Match 100.0%; Score 1327; DB 6; Length 3826;

```
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2559

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2560 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2619

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2679

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2680 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2739

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2740 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2799

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2800 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2859

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2860 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2919

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2920 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2979

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2980 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 3039

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3040 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 3099

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3100 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 3159

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3160 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 3219

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3220 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 3279

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3280 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 3339

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3340 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3399

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3400 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3459

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3460 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3519

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3520 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3579

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3580 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3639

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3640 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3699

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3700 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3759

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3760 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3819

Qy       1321 AAAAAAA 1327
              |||||||
Db       3820 AAAAAAA 3826
```

RESULT 2
AAK80624
ID AAK80624 standard; DNA; 5858 BP.
XX
AC AAK80624;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35436.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 35436; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 5858 BP; 1455 A; 1579 C; 1490 G; 1334 T; 0 U; 0 Other;

Query Match 98.6%; Score 1308; DB 4; Length 5858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1027 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1086

```
Qy        61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
Db      1087 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 1146

Qy       121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
Db      1147 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 1206

Qy       181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
Db      1207 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 1266

Qy       241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
Db      1267 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 1326

Qy       301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
Db      1327 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 1386

Qy       361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
Db      1387 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 1446

Qy       421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
Db      1447 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 1506

Qy       481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
Db      1507 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 1566

Qy       541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
Db      1567 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 1626

Qy       601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
Db      1627 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 1686

Qy       661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
Db      1687 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 1746

Qy       721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
Db      1747 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 1806

Qy       781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
Db      1807 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 1866

Qy       841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
Db      1867 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 1926

Qy       901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
Db      1927 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 1986

Qy       961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
Db      1987 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2046

Qy      1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
Db      2047 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 2106

Qy      1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
Db      2107 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 2166

Qy      1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
Db      2167 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 2226

Qy      1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
Db      2227 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 2286

Qy      1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
Db      2287 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 2334
```

RESULT 3
ABK12810
ID ABK12810 standard; DNA; 22893 BP.
XX
AC ABK12810;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human tumour suppressor CAR-1, BAC clone RP11-150F21 3' sequence.
XX
KW Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
KW colon cancer; stomach cancer; breast cancer; endometrial cancer;
KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;
KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
KW bacterila artificial chromosome; chromosome 1p31-1p36.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT misc_feature   5820
FT   /*tag= a
FT   /note= "Nucleotides 5821-6360 of the present sequence as
FT   reproduced in the specification are illegible or missing,
FT   nucleotide 5821 of the present sequence corresponds to
FT   nucletode 6361 of sequence as printed in the
FT   specification"
XX
PN WO200212285-A2.
XX
PD 14-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-US025269.
XX
PR 10-AUG-2000; 2000US-0225033P.
PR 23-AUG-2000; 2000US-0227560P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Killary A, Chandler D, Lott S;
XX
DR WPI; 2002-269088/31.
XX
PT New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
PT diagnosing cancer, for altering the phenotype of a tumor cell, for
PT treating cancers or as a diagnostic or prognostic indicator of cancer.
XX
PS Disclosure; Page 170-176; 185pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide being tumour suppressor, CAR-1. Also included are fragments
CC of the polynucleotide from 15-5000 nucleotides, fragments of the protein
CC from 10-50 amino acids, an expression cassette comprising the
CC polynucleotide under the control of a promoter operable in eukaryotic
CC cells, a method for suppressing growth of a cancer cells by contacting
CC the cells with the expression cassette (i.e. gene therapy), a cell
CC comprising the expression cassette, an anti-CAR-1 monoclonal or
CC polyclonal antibody, a hybridoma cell that produces the monoclonal
CC antibody, a method of diagnosing a cancer by assessing the expression of
CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,

```
CC   methods for altering the phenotype of a tumour cell, methods for treating
CC   a subject with cancer by administering the tumour suppressor CAR-1, or by
CC   administering a nucleic acid encoding the tumour suppressor CAR-1 and a
CC   promoter active in eukaryotic cells, where the promoter is operably
CC   linked to the region encoding the tumour suppressor, a non-human
CC   transgenic eukaryote lacking a functional CAR-1 gene, a non-human
CC   transgenic eukaryote that over-expresses CAR-1 as compared to a similar
CC   non-transgenic eukaryote, a method of screening a candidate substance for
CC   anti-tumour activity by contacting a cell lacking functional CAR-1
CC   polypeptide, with a candidate substance and determining the effect of the
CC   candidate substance on the cell, an anti-tumour composition produced by
CC   contacting a cell lacking functional CAR-1 polypeptide, with a candidate
CC   substance, determining the effect of the candidate substance on the cell,
CC   identifying a candidate inhibitor substance, and making a composition and
CC   an isolated and purified nucleic acid that hybridizes, under high
CC   stringency conditions, to a DNA segment comprising about 15-3826 bases of
CC   the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
CC   altering the phenotype of a tumour cell, for treating cancers (e.g.
CC   cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
CC   intestine, blood cells, colon, stomach, breast, endometrium, prostate,
CC   testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
CC   other tissues), and as a diagnostic or prognostic indicator of cancer.
CC   CAR-1 may also be used in screening compounds for activity in either
CC   stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
CC   effect of mutant CAR-1 molecule. The gene for CAR-1 is lcated on
CC   chromosome 1 (1p31-1p36). The present sequence is a BAC (bacterial
CC   artificial chromosome) containing part of the CAR-1 gene
XX
SQ   Sequence 22893 BP; 5720 A; 5989 C; 5664 G; 5415 T; 0 U; 105 Other;

  Query Match             98.6%;  Score 1308;  DB 6;  Length 22893;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1308;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13010 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 13069

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13070 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 13129

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13130 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 13189

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13190 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 13249

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13250 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 13309

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13310 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 13369

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13370 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 13429

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13430 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 13489

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13490 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 13549

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13550 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 13609

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13610 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 13669

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13670 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 13729

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13730 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 13789

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13790 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 13849

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13850 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 13909

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13910 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 13969

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13970 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 14029

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14030 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 14089

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14090 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 14149

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14150 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 14209

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14210 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 14269

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      14270 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 14317


RESULT 4
ADC37268
ID   ADC37268 standard; DNA; 3243 BP.
XX
AC   ADC37268;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 101.
XX
KW   Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW   cancer; infectious disease; bone disease; AIDS;
KW   neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW   Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW   Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003048202-A2.
XX
PD   12-JUN-2003.
XX
PF   03-DEC-2002; 2002WO-JP012644.
```

```
XX
PR   03-DEC-2001; 2001JP-00368692.
PR   05-DEC-2001; 2001US-0335829P.
PR   03-OCT-2002; 2002JP-00291302.
PR   04-OCT-2002; 2002US-0415769P.
XX
PA   (ASAH ) ASAHI KASEI KK.
XX
PI   Matsuda A,  Muramatsu S;
XX
DR   WPI; 2003-505282/47.
DR   P-PSDB; ADC37269.
XX
PT   New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT   useful for treating inflammation, autoimmune diseases, cancers,
PT   infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT   ischemic disorders.
XX
PS   Claim 4; SEQ ID NO 101; 938pp; English.
XX
CC   The present invention relates to novel proteins and their coding
CC   sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC   kappaB). The proteins and their coding sequences are useful for treating
CC   a disease associated with NF-kappaB activation, such as inflammation,
CC   autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC   neurodegenerative diseases, or ischaemic disorders.
XX
SQ   Sequence 3243 BP; 715 A; 1023 C; 863 G; 642 T; 0 U; 0 Other;

  Query Match             98.5%;  Score 1307;  DB 9;  Length 3243;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1307;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1937 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1996

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1997 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2056

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2057 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2116

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2117 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2176

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2177 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2236

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2237 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2296

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2297 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2356

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2357 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2416

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2417 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2476

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2477 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2536

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2537 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2596

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2597 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2656

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2657 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2716

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2717 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2776

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2777 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 2836

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2837 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 2896

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2897 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2956

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2957 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3016

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3017 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3076

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3077 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3136

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3137 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3196

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       3197 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 3243


RESULT 5
AAK80625
ID   AAK80625 standard; DNA; 5866 BP.
XX
AC   AAK80625;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35437.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
```

PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 35437; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)

CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 5866 BP; 1464 A; 1580 C; 1489 G; 1333 T; 0 U; 0 Other;

Query Match 94.7%; Score 1257; DB 4; Length 5866;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 1085

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 1145

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 1205

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 1265

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 1325

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 1385

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 1445

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 1505

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 1565

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 1625

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 1685

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 1745

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 1805

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 1865

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 1925

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 1985

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1986 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 2045

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2046 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 2105

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2106 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 2165

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2166 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 2225

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2226 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 2285

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2286 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 2333
```

RESULT 6
ABK12811
ID ABK12811 standard; DNA; 30676 BP.
XX
AC ABK12811;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human tumour suppressor CAR-1, BAC clone RP11-131M11.
XX
KW Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
KW colon cancer; stomach cancer; breast cancer; endometrial cancer;
KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;
KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
KW bacterila artificial chromosome; chromosome 1p31-1p36.
XX
OS Homo sapiens.
XX
PN WO200212285-A2.
XX
PD 14-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-US025269.
XX
PR 10-AUG-2000; 2000US-0225033P.
PR 23-AUG-2000; 2000US-0227560P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Killary A, Chandler D, Lott S;
XX
DR WPI; 2002-269088/31.
XX

```
PT   New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
PT   diagnosing cancer, for altering the phenotype of a tumor cell, for
PT   treating cancers or as a diagnostic or prognostic indicator of cancer.
XX
PS   Disclosure; Page 176-185; 185pp; English.
XX
CC   The invention relates to an isolated polynucleotide encoding a
CC   polypeptide being tumour suppressor, CAR-1. Also included are fragments
CC   of the polynucleotide from 15-5000 nucleotides, fragments of the protein
CC   from 10-50 amino acids, an expression cassette comprising the
CC   polynucleotide under the control of a promoter operable in eukaryotic
CC   cells, a method for suppressing growth of a cancer cells by contacting
CC   the cells with the expression cassette (i.e. gene therapy), a cell
CC   comprising the expression cassette, an anti-CAR-1 monoclonal or
CC   polyclonal antibody, a hybridoma cell that produces the monoclonal
CC   antibody, a method of diagnosing a cancer by assessing the expression of
CC   CAR-1 tumour suppressor in the cells of a tissue sample from a subject,
CC   methods for altering the phenotype of a tumour cell, methods for treating
CC   a subject with cancer by administering the tumour suppressor CAR-1, or by
CC   administering a nucleic acid encoding the tumour suppressor CAR-1 and a
CC   promoter active in eukaryotic cells, where the promoter is operably
CC   linked to the region encoding the tumour suppressor, a non-human
CC   transgenic eukaryote lacking a functional CAR-1 gene, a non-human
CC   transgenic eukaryote that over-expresses CAR-1 as compared to a similar
CC   non-transgenic eukaryote, a method of screening a candidate substance for
CC   anti-tumour activity by contacting a cell lacking functional CAR-1
CC   polypeptide, with a candidate substance and determining the effect of the
CC   candidate substance on the cell, an anti-tumour composition produced by
CC   contacting a cell lacking functional CAR-1 polypeptide, with a candidate
CC   substance, determining the effect of the candidate substance on the cell,
CC   identifying a candidate inhibitor substance, and making a composition and
CC   an isolated and purified nucleic acid that hybridizes, under high
CC   stringency conditions, to a DNA segment comprising about 15-3826 bases of
CC   the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
CC   altering the phenotype of a tumour cell, for treating cancers (e.g.
CC   cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
CC   intestine, blood cells, colon, stomach, breast, endometrium, prostate,
CC   testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
CC   other tissues), and as a diagnostic or prognostic indicator of cancer.
CC   CAR-1 may also be used in screening compounds for activity in either
CC   stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
CC   effect of mutant CAR-1 molecule. The gene for CAR-1 is located on
CC   chromosome 1 (1p31-1p36). The present sequence is a BAC (bacterial
CC   artificial chromosome) containing part of the CAR-1 gene
XX
SQ   Sequence 30676 BP; 7591 A; 7621 C; 7888 G; 6944 T; 0 U; 632 Other;

  Query Match             94.7%;  Score 1257;  DB 6;  Length 30676;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1307;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24440 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 24499

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24500 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 24559

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24560 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 24619

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24620 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 24679

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24680 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 24739

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      24740 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 24799

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24800 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 24859

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24860 GGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 24919

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24920 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 24979

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24980 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 25039

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25040 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 25099

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25100 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 25159

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25160 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 25219

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25220 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 25279

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25280 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 25339

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25340 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 25399

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25400 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 25459

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25460 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 25519

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25520 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 25579

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25580 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 25639

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25640 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 25699

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      25700 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 25747


RESULT 7
AAH14509
ID   AAH14509 standard; cDNA; 3436 BP.
XX
AC   AAH14509;
XX
```

DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12036.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12036; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 3436 BP; 731 A; 1126 C; 901 G; 678 T; 0 U; 0 Other;

Query Match 94.6%; Score 1256; DB 4; Length 3436;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1306; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2189

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2249

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2250 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2309

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2310 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2369

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2370 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2429

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2430 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2489

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2490 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2549

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2550 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2609

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2610 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 2669

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2670 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 2729

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2730 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 2789

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 2849

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 2909

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 2969

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 CATCTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3029

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3089

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3149

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3209

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3269

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3329

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3389
```

```
Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 1307
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTC 3436
```

RESULT 8
AAK80626
ID AAK80626 standard; DNA; 5858 BP.
XX
AC AAK80626;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35438.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.

```
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 35438; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 5858 BP; 1481 A; 1575 C; 1462 G; 1340 T; 0 U; 0 Other;

  Query Match           54.7%;  Score 726;  DB 4;  Length 5858;
  Best Local Similarity  99.4%;  Pred. No. 1.9e-267;
  Matches 1146;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        157 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1182 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 1241

Qy        217 GGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 276
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1242 GGTATACCCAAAGCTGATGCAAAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 1301

Qy        277 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1302 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 1361

Qy        337 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 1421

Qy        397 AAGCACT-GGGAGGGGCTGTTGGCTAGACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 455
              ||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       1422 AAGCACTGGGGAGGGGCTGTTGGCTAAACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 1481

Qy        456 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1482 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 1541

Qy        516 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1542 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 1601

Qy        576 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1602 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 1661

Qy        636 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1662 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 1721

Qy        696 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCC 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1722 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCC 1781

Qy        756 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1782 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 1841

Qy        816 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1842 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 1901

Qy        876 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGA 935
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1902 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAAAACTCGGGAGAGGGTGGA 1961

Qy        936 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1962 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 2021

Qy        996 GGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 1055
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2022 GGAGTGGGCTGCAAAACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 2081

Qy       1056 TGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 1115
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2082 TGGGGCCTAAATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 2141

Qy       1116 AGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 1175
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2142 AAAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 2201

Qy       1176 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2202 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 2261

Qy       1236 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2262 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 2321

Qy       1296 GTTACAGATGTCA 1308
              |||||||||||||
Db       2322 GTTACAGATGTCA 2334


RESULT 9
ABK12808
ID   ABK12808 standard; DNA; 30625 BP.
XX
AC   ABK12808;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human tumour suppressor CAR-1, BAC clone 392H05 3' sequence.
XX
KW   Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
KW   gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
```

KW colon cancer; stomach cancer; breast cancer; endometrial cancer;
KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;
KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
KW bacterila artificial chromosome; chromosome 1p31-1p36.
XX
OS Homo sapiens.
XX
PN WO200212285-A2.
XX
PD 14-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-US025269.
XX
PR 10-AUG-2000; 2000US-0225033P.
PR 23-AUG-2000; 2000US-0227560P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Killary A, Chandler D, Lott S;
XX
DR WPI; 2002-269088/31.
XX
PT New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
PT diagnosing cancer, for altering the phenotype of a tumor cell, for
PT treating cancers or as a diagnostic or prognostic indicator of cancer.
XX
PS Disclosure; Page 149-157; 185pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide being tumour suppressor, CAR-1. Also included are fragments
CC of the polynucleotide from 15-5000 nucleotides, fragments of the protein
CC from 10-50 amino acids, an expression cassette comprising the
CC polynucleotide under the control of a promoter operable in eukaryotic
CC cells, a method for suppressing growth of a cancer cells by contacting
CC the cells with the expression cassette (i.e. gene therapy), a cell
CC comprising the expression cassette, an anti-CAR-1 monoclonal or
CC polyclonal antibody, a hybridoma cell that produces the monoclonal
CC antibody, a method of diagnosing a cancer by assessing the expression of
CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,
CC methods for altering the phenotype of a tumour cell, methods for treating
CC a subject with cancer by administering the tumour suppressor CAR-1, or by
CC administering a nucleic acid encoding the tumour suppressor CAR-1 and a
CC promoter active in eukaryotic cells, where the promoter is operably
CC linked to the region encoding the tumour suppressor, a non-human
CC transgenic eukaryote lacking a functional CAR-1 gene, a non-human
CC transgenic eukaryote that over-expresses CAR-1 as compared to a similar
CC non-transgenic eukaryote, a method of screening a candidate substance for
CC anti-tumour activity by contacting a cell lacking functional CAR-1
CC polypeptide, with a candidate substance and determining the effect of the
CC candidate substance on the cell, an anti-tumour composition produced by
CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate
CC substance, determining the effect of the candidate substance on the cell,
CC identifying a candidate inhibitor substance, and making a composition and
CC an isolated and purified nucleic acid that hybridizes, under high
CC stringency conditions, to a DNA segment comprising about 15-3826 bases of
CC the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
CC altering the phenotype of a tumour cell, for treating cancers (e.g.
CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,
CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
CC other tissues), and as a diagnostic or prognostic indicator of cancer.
CC CAR-1 may also be used in screening compounds for activity in either
CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
CC effect of mutant CAR-1 molecule. The gene for CAR-1 is lcated on
CC chromosome 1 (1p31-1p36). The present sequence is a BAC (bacterial
CC artificial chromosome) containing part of the CAR-1 gene
XX
SQ Sequence 30625 BP; 8084 A; 7712 C; 7487 G; 7314 T; 0 U; 28 Other;

Query Match 54.7%; Score 726; DB 6; Length 30625;
Best Local Similarity 99.4%; Pred. No. 1.4e-267;
Matches 1146; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy      157 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 216
Db    23131 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 23190

Qy      217 GGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 276
Db    23191 GGTATACCCAAAGCTGATGCAAAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 23250

Qy      277 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 336
Db    23251 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 23310

Qy      337 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 396
Db    23311 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 23370

Qy      397 AAGCACT-GGGAGGGGCTGTTGGCTAGACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 455
Db    23371 AAGCACTGGGGAGGGGCTGTTGGCTAAACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 23430

Qy      456 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 515
Db    23431 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 23490

Qy      516 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 575
Db    23491 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 23550

Qy      576 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 635
Db    23551 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 23610

Qy      636 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 695
Db    23611 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 23670

Qy      696 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCC 755
Db    23671 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCC 23730

Qy      756 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 815
Db    23731 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 23790

Qy      816 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 875
Db    23791 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 23850

Qy      876 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGA 935
Db    23851 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAAAACTCGGGAGAGGGTGGA 23910

Qy      936 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 995
Db    23911 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 23970

Qy      996 GGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 1055
Db    23971 GGAGTGGGCTGCAAAACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 24030

Qy     1056 TGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 1115
Db    24031 TGGGGCCTAAATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 24090

Qy     1116 AGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 1175
Db    24091 AAAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 24150

Qy     1176 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 1235
Db    24151 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 24210

Qy     1236 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 1295
```

```
Db      24211 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 24270

Qy       1296 GTTACAGATGTCA 1308
              |||||||||||||
Db      24271 GTTACAGATGTCA 24283


RESULT 10
ADC37454
ID   ADC37454 standard; DNA; 2207 BP.
XX
AC   ADC37454;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 287.
XX
KW   Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW   cancer; infectious disease; bone disease; AIDS;
KW   neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW   Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW   Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003048202-A2.
XX
PD   12-JUN-2003.
XX
PF   03-DEC-2002; 2002WO-JP012644.
XX
PR   03-DEC-2001; 2001JP-00368692.
PR   05-DEC-2001; 2001US-0335829P.
PR   03-OCT-2002; 2002JP-00291302.
PR   04-OCT-2002; 2002US-0415769P.
XX
PA   (ASAH ) ASAHI KASEI KK.
XX
PI   Matsuda A,  Muramatsu S;
XX
DR   WPI; 2003-505282/47.
DR   P-PSDB; ADC37455.
XX
PT   New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT   useful for treating inflammation, autoimmune diseases, cancers,
PT   infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT   ischemic disorders.
XX
PS   Claim 4; SEQ ID NO 287; 938pp; English.
XX
CC   The present invention relates to novel proteins and their coding
CC   sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC   kappaB). The proteins and their coding sequences are useful for treating
CC   a disease associated with NF-kappaB activation, such as inflammation,
CC   autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC   neurodegenerative diseases, or ischaemic disorders.
XX
SQ   Sequence 2207 BP; 414 A; 784 C; 618 G; 391 T; 0 U; 0 Other;

  Query Match           34.1%;  Score 452;  DB 9;  Length 2207;
  Best Local Similarity  99.8%;  Pred. No. 7.4e-163;
  Matches  502;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        801 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1705 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 1764

Qy        861 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC 920
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1765 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTCTTCTGCCACAGCATCCAATCGTGAAGAAC 1824

Qy        921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1825 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 1884

Qy        981 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1885 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1944

Qy       1041 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1945 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 2004

Qy       1101 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2005 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 2064

Qy       1161 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2065 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 2124

Qy       1221 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2125 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 2184

Qy       1281 TTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||
Db       2185 TTGTATTAAATTAAAGTTACAGA 2207


RESULT 11
ADC37264
ID   ADC37264 standard; DNA; 2207 BP.
XX
AC   ADC37264;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 97.
XX
KW   Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW   cancer; infectious disease; bone disease; AIDS;
KW   neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW   Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW   Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003048202-A2.
XX
PD   12-JUN-2003.
XX
PF   03-DEC-2002; 2002WO-JP012644.
XX
PR   03-DEC-2001; 2001JP-00368692.
PR   05-DEC-2001; 2001US-0335829P.
PR   03-OCT-2002; 2002JP-00291302.
PR   04-OCT-2002; 2002US-0415769P.
XX
PA   (ASAH ) ASAHI KASEI KK.
XX
PI   Matsuda A,  Muramatsu S;
XX
DR   WPI; 2003-505282/47.
DR   P-PSDB; ADC37265.
XX
PT   New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT   useful for treating inflammation, autoimmune diseases, cancers,
PT   infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT   ischemic disorders.
XX
PS   Claim 4; SEQ ID NO 97; 938pp; English.
XX
CC   The present invention relates to novel proteins and their coding
CC   sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
```

CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 2207 BP; 414 A; 784 C; 618 G; 391 T; 0 U; 0 Other;

Query Match 34.1%; Score 452; DB 9; Length 2207;
Best Local Similarity 99.8%; Pred. No. 7.4e-163;
Matches 502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        801 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1705 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 1764

Qy        861 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC 920
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1765 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTCTTCTGCCACAGCATCCAATCGTGAAGAAC 1824

Qy        921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1825 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 1884

Qy        981 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1885 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1944

Qy       1041 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1945 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 2004

Qy       1101 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2005 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 2064

Qy       1161 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2065 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 2124

Qy       1221 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2125 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 2184

Qy       1281 TTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||
Db       2185 TTGTATTAAATTAAAGTTACAGA 2207
```

RESULT 12
ADC37266
ID ADC37266 standard; DNA; 2246 BP.
XX
AC ADC37266;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 99.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAHI KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PSDB; ADC37267.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 99; 938pp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 2246 BP; 418 A; 805 C; 623 G; 400 T; 0 U; 0 Other;

Query Match 34.1%; Score 452; DB 9; Length 2246;
Best Local Similarity 99.8%; Pred. No. 7.4e-163;
Matches 502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        801 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 CAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTT 1803

Qy        861 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC 920
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1804 AGGGTGGGTGGAGCAAGCCCCACCTGCGTTCTTCTGCCACAGCATCCAATCGTGAAGAAC 1863

Qy        921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 TCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCA 1923

Qy        981 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 GAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGAC 1983

Qy       1041 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 TTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCAT 2043

Qy       1101 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 CCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCC 2103

Qy       1161 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2104 GGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATC 2163

Qy       1221 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2164 TGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAAT 2223

Qy       1281 TTGTATTAAATTAAAGTTACAGA 1303
              |||||||||||||||||||||||
Db       2224 TTGTATTAAATTAAAGTTACAGA 2246
```

RESULT 13
AAH11818/c
ID AAH11818 standard; cDNA; 557 BP.
XX
AC AAH11818;

XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (3'-primer) SEQ ID NO:8653.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 3; SEQ ID NO 8653; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 557 BP; 142 A; 133 C; 139 G; 136 T; 0 U; 7 Other;

Query Match 31.9%; Score 423; DB 4; Length 557;
Best Local Similarity 100.0%; Pred. No. 1.1e-151;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       885 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 944
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       423 TGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 364

Qy       945 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 1004
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       363 TAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGC 304

Qy      1005 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 1064
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       303 TGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTA 244

Qy      1065 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 1124
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       243 GATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 184

Qy      1125 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 1184
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       183 CACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAAT 124

Qy      1185 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 1244
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       123 TGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTA 64

Qy      1245 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 1304
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGAT 4

Qy      1305 GTC 1307
            |||
Db          3 GTC 1
```

RESULT 14
ABN43449
ID ABN43449 standard; DNA; 60 BP.
XX
AC ABN43449;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:16197.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB001903.
XX
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
PS Example 1; SEQ ID NO 16197; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini

```
CC   libraries to detect transcripts of a sub-transcriptome under a particular
CC   biological or pathological state, and so allowing the detection of tissue
CC   - and pathology-specific genes such as those genes only expressed in
CC   specific tissue under a specific pathological condition; to detect
CC   developmental specific genes; and to detect RNA transcripts and splice
CC   variants of a transcriptome of a patient suffering from a particular
CC   disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC   rats, humans and mice, which are used in the exemplification of the
CC   present invention. N.B. The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 60 BP; 17 A; 21 C; 11 G; 11 T; 0 U; 0 Other;

  Query Match             4.5%;  Score 60;  DB 6;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-13;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1169 CTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCAC 1228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCAC 60


RESULT 15
ABN63808
ID   ABN63808 standard; cDNA; 614 BP.
XX
AC   ABN63808;
XX
DT   28-JUN-2002  (first entry)
XX
DE   Human cancer related polynucleotide SEQ ID NO 3775.
XX
KW   Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW   gene therapy; cancer; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200214500-A2.
XX
PD   21-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-US025840.
XX
PR   16-AUG-2000; 2000US-0226326P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J,  Garcia PD,  Sudduth-Klinger J,  Reinhard C,  Randazzo F;
PI   Lamson G,  Scott EM,  Zhang G,  Kassam A,  Pot D,  Labat I;
XX
DR   WPI; 2002-241905/29.
XX
PT   New nucleic acid for producing a polypeptide, detecting differentially
PT   expressed genes correlated with a cancerous state of a mammalian cell,
PT   and inhibiting tumor growth.
XX
PS   Claim 1; SEQ ID NO 3775; 883pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC   with cytostatic activity. The polynucleotide is used to produce a
CC   polypeptide, to detect differentially expressed genes correlated with a
CC   cancerous state of a mammalian cell and to inhibit tumour growth. The
CC   polynucleotide is used as a probe in mapping and tissue profiling. The
CC   encoded polypeptide and antibodies to the polypeptide can also be used
CC   for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC   gene therapy. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 614 BP; 187 A; 127 C; 137 G; 163 T; 0 U; 0 Other;

  Query Match             2.2%;  Score 29;  DB 6;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||||||
Db        585 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 613


RESULT 16
AAC59529
ID   AAC59529 standard; cDNA; 470 BP.
XX
AC   AAC59529;
XX
DT   15-FEB-2001  (first entry)
XX
DE   Human secreted protein cDNA sequence #23.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055352-A2.
XX
PD   21-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US006044.
XX
PR   12-MAR-1999;   99US-0124099P.
PR   03-DEC-1999;   99US-0168661P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-602124/57.
DR   P-PSDB; AAB34238.
XX
PT   Novel human secreted proteins useful for diagnosis, prevention and
PT   treatment of disorders including neurological, cell proliferative,
PT   cardiovascular, autoimmune and inflammatory disorders and microbial
PT   infections.
XX
PS   Claim 1; Page 331; 383pp; English.
XX
CC   The invention relates to the isolation of genes AAC59507-C59556 encoding
CC   50 human secreted proteins AAB34218-B34264) The genes can be used to
CC   generate fusion proteins by linking to the gene for the human
CC   immunoglobulin G Fc portion (AAC59498) for increasing the stability of
CC   the fusion protein as compared to the human protein only. The genes and
CC   proteins are useful for preventing, ameliorating or treating medical
CC   conditions, e.g. by protein or gene therapy. The genes are isolated from
CC   a range of human tissues disclosed in the specification. The nucleic
CC   acids, proteins, antibodies and (ant)agonists are useful in the
CC   diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC   immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections
XX
SQ   Sequence 470 BP; 165 A; 90 C; 66 G; 149 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 27;  DB 3;  Length 470;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db        410 AGATGTCAAAAAAAAAAAAAAAAAAAA 436


RESULT 17
ACF62741
ID   ACF62741 standard; DNA; 189013 BP.
XX
AC   ACF62741;
XX
DT   08-OCT-2003  (first entry)
XX
DE   Cancer based on CYP3A5 related polynucleotide SEQ ID NO:669.
XX
KW   Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KW   cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;
KW   cytostatic; gene; ds.
XX
OS   Unidentified.
XX
PN   WO2003013534-A2.
XX
PD   20-FEB-2003.
XX
PF   23-JUL-2002; 2002WO-EP008219.
XX
PR   23-JUL-2001; 2001EP-00117608.
PR   24-MAY-2002; 2002EP-00011710.
XX
PA   (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI   Heinrich G,  Kerb R;
XX
DR   WPI; 2003-268144/26.
XX
PT   New use of irinotecan for preparation of compositions for treating cancer
PT   in subject having genome with variant allele comprising cytochrome p450,
PT   subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX
PS   Disclosure; SEQ ID NO 669; 86pp; English.
XX
CC   The present invention describes the use of irinotecan (I) or its
CC   derivative for the preparation of a pharmaceutical composition for
CC   treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC   cancer, or malignant glioma in a subject having a genome with a variant
CC   allele which comprises a cytochrome p450, subfamily IIIA (nifedipine
CC   oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
CC   cytostatic activity. The therapeutic applications of (I) is improved,
CC   since it is possible to individually treat a subject with an appropriate
CC   dosage and/or an appropriate derivative of (I). Therefore, undesirable,
CC   harmful or toxic effects are efficiently avoided. Unnecessary and
CC   potentially harmful treatment of those subjects who do not respond to the
CC   treatment with substances (nonresponders), as well as the development of
CC   drug resistances due to suboptimal drug dosing can be avoided. ACF62200
CC   to ACF62751 and ABM34912 to ABM35013 represent sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 27;  DB 7;  Length 189013;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db     123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963


RESULT 18
ADB20856
ID   ADB20856 standard; DNA; 189013 BP.
XX
AC   ADB20856;
XX
DT   20-NOV-2003  (first entry)
XX
DE   MRP1 based cancer related nucleic acid SEQ ID NO:669.
XX
KW   irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW   lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW   variant allele; multidrug resistance protein 1; MRP1; cytostatic; gene;
KW   ds.
XX
OS   Unidentified.
XX
PN   WO2003013533-A2.
XX
PD   20-FEB-2003.
XX
PF   23-JUL-2002; 2002WO-EP008200.
XX
PR   23-JUL-2001; 2001EP-00117608.
PR   24-MAY-2002; 2002EP-00011710.
XX
PA   (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI   Heinrich G,  Kerb R;
XX
DR   WPI; 2003-354397/33.
XX
PT   Use of irinotecan or its derivative for preparation of a pharmaceutical
PT   composition for treating cancer in a subject having a genome with a
PT   variant allele comprising a multidrug resistance protein 1
PT   polynucleotide.
XX
PS   Disclosure; SEQ ID NO 669; 100pp; English.
XX
CC   The present invention describes a method for the use of irinotecan (I) or
CC   its derivative for the preparation of a pharmaceutical composition for
CC   treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC   cancer, or malignant glioma in a subject having a genome with a variant
CC   allele which comprises a multidrug resistance protein 1 (MRP1)
CC   polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
CC   can be used for the preparation of a pharmaceutical composition for
CC   treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC   cancer, or malignant glioma in a subject, where the subject is a human
CC   (preferably African or Asian) or a mouse. The present sequence represents
CC   a sequence which is used in the exemplification of the present invention.
XX
SQ   Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 27;  DB 7;  Length 189013;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||
Db     123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963


RESULT 19
ADB87945
ID   ADB87945 standard; DNA; 189013 BP.
XX
AC   ADB87945;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human UGT1A1 gene sequence SEQ ID NO:669.
XX
KW   irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;
KW   colorectal cancer; cervical cancer; gastric cancer; lung cancer;
```

KW ovarian cancer; pancreatic cancer; malignant glioma;
KW uridine diphosphate glycosyltransferase1 member A1; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003013536-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008217.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-289896/28.
XX
PT Use of irinotecan to treat cancer patient by determining if patient has
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.
XX
PS Disclosure; SEQ ID NO 669; 107pp; English.
XX
CC The invention relates to the novel use of irinotecan to treat a patient
CC suffering from cancer. This involves determining if the patient has one
CC or more variant alleles of the UGT1A1 gene, and if the patient has one or
CC more of such variant alleles, irinotecan is administered in an increased
CC or decreased amount in comparison to the amount that is administered
CC without regard to the patient's alleles in the UGT1A1 gene. The invention
CC has cytostatic activity. A composition of the invention acts as a
CC topoisomerase I inhibitor. The method is useful for treating a patient,
CC an animal e.g. mouse or a human, preferably African or Asian, suffering
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
CC pancreatic cancer or malignant glioma. The present sequence is udes in
CC the exemplification of the invention.
XX
SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 9; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||||
Db 123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963

RESULT 20
ADB96928
ID ADB96928 standard; DNA; 189013 BP.
XX
AC ADB96928;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human MDR1 related DNA sequence SEQ ID NO:669.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDR1; cytostatic; human; Cyp3A5; MRP1; MDR1;
KW TOP1; ds.
XX
OS Homo sapiens.
XX
PN WO2003013537-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008218.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268145/26.
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 669; 130pp; English.
XX
CC The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytostatic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX
SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 9; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||||
Db 123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963

RESULT 21
ADB92119
ID ADB92119 standard; DNA; 189013 BP.
XX
AC ADB92119;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human MDR1 related DNA sequence SEQ ID NO:669.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOP1; ds.
XX
OS Homo sapiens.
XX
PN WO2003013535-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008220.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-342400/32.
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX

PS Disclosure; SEQ ID NO 669; 104pp; English.
XX
CC The invention relates ro a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 9; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 AGATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||||
Db 123937 AGATGTCAAAAAAAAAAAAAAAAAAAA 123963

RESULT 22
ABQ60500
ID ABQ60500 standard; cDNA; 273 BP.
XX
AC ABQ60500;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:4195.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 273 BP; 106 A; 29 C; 47 G; 91 T; 0 U; 0 Other;

Query Match 2.0%; Score 26; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 245 GATGTCAAAAAAAAAAAAAAAAAAAA 270

RESULT 23
ABX41090/c
ID ABX41090 standard; cDNA; 393 BP.
XX
AC ABX41090;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #6255.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 6255; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or

```
CC   for genetically improving cattle. The present sequence is one of the
CC   15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC   present sequence was not shown in the specification but was obtained in
CC   electronic format from the USPTO web site:
CC   seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ   Sequence 393 BP; 107 A; 75 C; 96 G; 115 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 26;  DB 7;  Length 393;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||
Db         49 GATGTCAAAAAAAAAAAAAAAAAAAA 24


RESULT 24
ABX48150/c
ID   ABX48150 standard; cDNA; 411 BP.
XX
AC   ABX48150;
XX
DT   21-FEB-2003  (first entry)
XX
DE   Bovine EST associated with lactation/muscle/fat deposition #13315.
XX
KW   Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW   muscle deposition; fat deposition; genome mapping; gene identification;
KW   gene analysis; cattle breeding.
XX
OS   Bos Taurus.
XX
PN   US2002137139-A1.
XX
PD   26-SEP-2002.
XX
PF   24-SEP-2001; 2001US-00960352.
XX
PR   12-JAN-1999;   99US-0115707P.
PR   11-JAN-2000; 2000US-00480902.
XX
PA   (BYAT/) BYATT J C.
PA   (MATH/) MATHIALAGAN N.
PA   (TAON/) TAO N.
PA   (WARR/) WARREN W C.
XX
PI   Byatt JC,  Mathialagan N,  Tao N,  Warren WC;
XX
DR   WPI; 2003-110599/10.
XX
PT   New nucleic acid associated with lactation, and muscle and fat
PT   deposition, useful for genome mapping, gene identification and analysis,
PT   cattle breeding, or for genetically improving cattle.
XX
PS   Claim 2; SEQ ID NO 13315; 245pp; English.
XX
CC   The invention relates to a purified nucleic acid molecule associated with
CC   lactation or muscle and fat deposition (designated LMFD), derived from
CC   cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC   nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC   appearing as ABX34836-ABX49947, or complements of them. Also included are
CC   ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC   acid linked to a promoter and a 3' non- translated sequence that
CC   functions in the cell to cause termination of transcription and addition
CC   of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC   (2) determining a level or pattern of a molecule in a bovine cell or
CC   tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC   of the 15112 nucleic acid sequences or its complement or fragment) with a
CC   complementary nucleic acid molecule obtained from the bovine cell or
CC   tissue, where hybridisation between the marker nucleic acid and the
CC   complementary nucleic acid permits the detection of the molecule; and (b)
CC   detecting the level or pattern of the complementary nucleic acid, where
CC   the detection of the complementary nucleic acid is predictive of the
CC   level or pattern of the molecule. The LMFD nucleic acid is used for
CC   determining a level or pattern of a molecule in a bovine cell or tissue.
CC   It is useful for genome mapping, gene identification and analysis, cattle
CC   breeding, preparation of constructs for use in cattle gene expression, or
CC   for genetically improving cattle. The present sequence is one of the
CC   15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC   present sequence was not shown in the specification but was obtained in
CC   electronic format from the USPTO web site:
CC   seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ   Sequence 411 BP; 110 A; 80 C; 99 G; 122 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 26;  DB 7;  Length 411;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||
Db         48 GATGTCAAAAAAAAAAAAAAAAAAAA 23


RESULT 25
AAL24218/c
ID   AAL24218 standard; cDNA; 659 BP.
XX
AC   AAL24218;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Human breast cancer expressed polynucleotide 16675.
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151628-A2.
XX
PD   19-JUL-2001.
XX
PF   10-JAN-2001; 2001WO-US000798.
XX
PR   14-JAN-2000; 2000US-0176077P.
PR   14-MAR-2000; 2000US-0189167P.
PR   24-MAR-2000; 2000US-0192099P.
PR   29-MAR-2000; 2000US-0193480P.
PR   15-MAY-2000; 2000US-0205230P.
PR   09-JUN-2000; 2000US-0211315P.
PR   25-JUL-2000; 2000US-0220534P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lillie J,  Xu Y,  Wang Y,  Steinmann K;
XX
DR   WPI; 2001-451856/48.
XX
PT   New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS   Claim 1; Page 3055; 3695pp; English.
XX
CC   The invention relates to human breast cancer expressed polynucleotides
CC   (AAL07544-AAL26789) and methods of assessing whether a patient is
CC   afflicted with breast cancer by examining the correlation between the
CC   expression of certain markers and the cancerous state of breast cells.
CC   The polynucleotides and encoded polypeptides are potential markers for
CC   detecting, diagnosing, monitoring, characterising treating and
CC   potentially preventing breast cancer. The polynucleotides and encoded
CC   polypeptides are also useful for isolating compounds with cytostatic
CC   activity
XX
SQ   Sequence 659 BP; 184 A; 181 C; 110 G; 181 T; 0 U; 3 Other;

  Query Match             2.0%;  Score 26;  DB 4;  Length 659;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||
Db        599 GATGTCAAAAAAAAAAAAAAAAAAAA 574


RESULT 26
AAQ35208
ID   AAQ35208 standard; cDNA; 729 BP.
XX
AC   AAQ35208;
XX
DT   06-JUN-1993  (first entry)
XX
DE   Sequence encoding adenine phosphoribosyltransferase (apt).
XX
KW   Adenine phosphoribosyltransferase; probe; APRT activity; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             18. .569
FT                   /*tag=  a
XX
PN   CA2069262-A.
XX
PD   24-NOV-1992.
XX
PF   22-MAY-1992;   92CA-02069262.
XX
PR   23-MAY-1991;   91GB-00011126.
XX
PA   (UYWA-) UNIV WATERLOO.
XX
PI   Moffatt B;
XX
DR   WPI; 1993-053310/07.
DR   P-PSDB; AAR30826.
XX
PT   Purified c DNA for apt. gene in plants - used for analysis of adenine
PT   phosphoribosyl-transferase activity and function in plant development.
XX
PS   Claim 3; Fig 1; 27pp; English.
XX
CC   The apt cDNA of A. thaliana is described as a full-length clone. However,
CC   the inventors cannot unequivocally rule out that there are 5' sequences
CC   missing from this cDNA. The cDNA predicts a protein molecular weight of
CC   27,140
XX
SQ   Sequence 729 BP; 204 A; 123 C; 188 G; 214 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 26;  DB 2;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||
Db        703 GATGTCAAAAAAAAAAAAAAAAAAAA 728


RESULT 27
AAS02396
ID   AAS02396 standard; cDNA; 2295 BP.
XX
AC   AAS02396;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human secreted protein, cDNA #2.
XX
KW   Human; secreted protein; immunogen; antibody; diagnosis;
KW   rheumatoid arthritis; hyperproliferative disorder; neoplasm;
KW   cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW   angiogenesis; Alzheimer's disease; bacterial infection; viral infection;
KW   fungal infection; corneal infection; wound healing; cell culture;
KW   epithelial cell proliferation; skin ageing; transplantation;
KW   tissue regeneration; chemotaxis; food additive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200123546-A1.
XX
PD   05-APR-2001.
XX
PF   26-SEP-2000; 2000WO-US026323.
XX
PR   27-SEP-1999;   99US-0155805P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben S,  Komatsoulis GA;
XX
DR   WPI; 2001-266150/27.
DR   P-PSDB; AAU01723, AAU01759, AAU01760.
XX
PT   Nucleic acids encoding 37 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease
PT   and diabetic retinopathy.
XX
PS   Disclosure; Page 411; 494pp; English.
XX
CC   The sequence encodes a human secreted protein of the invention. The
CC   polynucleotides, polypeptides and antibodies raised against them are used
CC   to prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC   polynucleotides and antibodies are also used in diagnosing a pathological
CC   condition or susceptibility to a pathological condition. The antibodies
CC   can also be used in alleviating symptoms associated with the disorders
CC   and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC   immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC   include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC   disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC   e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC   angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC   infections caused by bacteria, viruses and fungi and ocular disorders
CC   e.g. corneal infection. The polypeptides can also be used to aid wound
CC   healing and epithelial cell proliferation, to prevent skin aging due to
CC   sunburn, to maintain organs before transplantation, for supporting cell
CC   culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities
XX
SQ   Sequence 2295 BP; 794 A; 393 C; 449 G; 659 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 26;  DB 4;  Length 2295;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAA 1326
             ||||||||||||||||||||||||||
Db       2270 AGATGTCAAAAAAAAAAAAAAAAAAA 2295


RESULT 28
ABX76356
ID   ABX76356 standard; DNA; 4702 BP.
XX
AC   ABX76356;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polynucleotide #220.
XX
KW   Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
```

```
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO200286443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US012476.
XX
PR   18-APR-2001; 2001US-0284770P.
PR   10-MAY-2001; 2001US-0290492P.
PR   09-NOV-2001; 2001US-0339245P.
PR   13-NOV-2001; 2001US-0350666P.
PR   29-NOV-2001; 2001US-0334370P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Aziz N,  Murray R;
XX
DR   WPI; 2003-093161/08.
DR   P-PSDB; ABU56627.
XX
PT   Detecting a lung cancer-associated transcript in a cell from a patient
PT   for treating lung cancer, by contacting a biological sample from the
PT   patient with a polynucleotide that exhibits increased or decreased
PT   expression in lung cancer.
XX
PS   Claim 22; Page 355-356; 453pp; English.
XX
CC   The invention relates to a method for detecting a lung cancer-associated
CC   transcript in a cell from a patient, comprising contacting a biological
CC   sample from the patient with a polynucleotide that selectively hybridises
CC   to a sequence that is at least 80 % identical to a gene that exhibits
CC   increased or decreased expression in lung cancer samples. Lung cancer-
CC   associated polynucleotides and polypeptides are used for identifying a
CC   compound that modulates a lung cancer-associated polypeptide, for
CC   inhibiting proliferation of a lung cancer-associated cell to treat lung
CC   cancer in a patient and for treating a mammal having lung cancer by
CC   administering a modulatory compound identified. The methods are useful
CC   for treating lung cancer, such as small cell lung cancer, non-small cell
CC   lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC   emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC   hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC   bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC   for diagnostic purposes and as targets for screening for therapeutic
CC   compounds that modulate lung cancer, such as antibodies. Sequences
CC   ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC   invention
XX
SQ   Sequence 4702 BP; 1159 A; 1218 C; 1217 G; 1108 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 26;  DB 7;  Length 4702;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||||
Db       4660 GATGTCAAAAAAAAAAAAAAAAAAAA 4685


RESULT 29
AAF22305_07
Continuation (8 of 11) of AAF22305 from base 700001 (Arabidopsis thaliana chromosome 4 c
WP Sequence split into 11 fragments  LOCUS AAF22305 Accession Aaf22305
WP     Fragment Name      Begin     End
WP     AAF22305_00            1    110000
WP     AAF22305_01       100001    210000
WP     AAF22305_02       200001    310000
WP     AAF22305_03       300001    410000
WP     AAF22305_04       400001    510000
WP     AAF22305_05       500001    610000
WP     AAF22305_06       600001    710000
WP     AAF22305_07       700001    810000
WP     AAF22305_08       800001    910000
WP     AAF22305_09       900001   1010000
WP     AAF22305_10      1000001   1082138

  Query Match             2.0%;  Score 26;  DB 3;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAA 1326
             ||||||||||||||||||||||||||
Db      16814 AGATGTCAAAAAAAAAAAAAAAAAAA 16839


RESULT 30
ABL86275
ID   ABL86275 standard; cDNA; 184 BP.
XX
AC   ABL86275;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human ovarian cancer related cDNA clone SEQ ID NO:9253.
XX
KW   Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200192581-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US017756.
XX
PR   26-MAY-2000; 2000US-0207484P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Harlocker SL,  Jones R;
XX
DR   WPI; 2002-122075/16.
XX
PT   Composition for therapy and diagnosis of ovarian cancer comprising
PT   polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT   polypeptide, antibody specific to polypeptide or T cell expressing
PT   polypeptide.
XX
PS   Claim 1; SEQ ID NO 9253; 489pp; English.
XX
CC   The present invention describes a composition (I) comprising: carriers
CC   and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC   polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC   from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC   (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC   or antigen presenting cells that express (II). (I) has cytostatic
CC   activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC   detecting ovarian cancer in a patient's biological sample preferably
CC   serum or ovarian tissue. The method comprises contacting a biological
CC   sample from a patient with (IV), detecting the amount of polynucleotide
CC   hybridising to (IV) and comparing the amount to a predetermined cutoff
CC   value and thereby detecting ovarian cancer in the patient, where the
CC   amount of polynucleotide hybridising to (IV) is detected preferably by
CC   polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC   useful for stimulating and/or expanding T cells specific for an ovarian
CC   tumour protein comprising contacting T cells with (III) or (II). (III) is
CC   useful in design and preparation of ribozyme molecules for inhibiting
CC   expression of the tumour polypeptides and proteins in tumour cells; and
CC   to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC   library using well known techniques
```

XX
SQ Sequence 184 BP; 76 A; 30 C; 50 G; 28 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 6; Length 184;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 141 ATGTCAAAAAAAAAAAAAAAAAAAA 165

RESULT 31
ABL65710/c
ID ABL65710 standard; DNA; 191 BP.
XX
AC ABL65710;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4047.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 4047; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 191 BP; 52 A; 38 C; 23 G; 78 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 6; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 30 ATGTCAAAAAAAAAAAAAAAAAAAA 6

RESULT 32
ABL64256/c
ID ABL64256 standard; DNA; 191 BP.
XX
AC ABL64256;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2593.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.

PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 2593; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 191 BP; 52 A; 38 C; 23 G; 78 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 6; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 30 ATGTCAAAAAAAAAAAAAAAAAAAA 6

RESULT 33
ABN94027/c
ID ABN94027 standard; DNA; 191 BP.
XX
AC ABN94027;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #525 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 525; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 191 BP; 52 A; 38 C; 23 G; 78 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 6; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 30 ATGTCAAAAAAAAAAAAAAAAAAAA 6

RESULT 34

```
ABL87801
ID   ABL87801 standard; cDNA; 207 BP.
XX
AC   ABL87801;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human ovarian cancer related cDNA clone SEQ ID NO:10779.
XX
KW   Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200192581-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US017756.
XX
PR   26-MAY-2000; 2000US-0207484P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Harlocker SL,  Jones R;
XX
DR   WPI; 2002-122075/16.
XX
PT   Composition for therapy and diagnosis of ovarian cancer comprising
PT   polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT   polypeptide, antibody specific to polypeptide or T cell expressing
PT   polypeptide.
XX
PS   Claim 1; SEQ ID NO 10779; 489pp; English.
XX
CC   The present invention describes a composition (I) comprising: carriers
CC   and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC   polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC   from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC   (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC   or antigen presenting cells that express (II). (I) has cytostatic
CC   activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC   detecting ovarian cancer in a patient's biological sample preferably
CC   serum or ovarian tissue. The method comprises contacting a biological
CC   sample from a patient with (IV), detecting the amount of polynucleotide
CC   hybridising to (IV) and comparing the amount to a predetermined cutoff
CC   value and thereby detecting ovarian cancer in the patient, where the
CC   amount of polynucleotide hybridising to (IV) is detected preferably by
CC   polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC   useful for stimulating and/or expanding T cells specific for an ovarian
CC   tumour protein comprising contacting T cells with (III) or (II). (III) is
CC   useful in design and preparation of ribozyme molecules for inhibiting
CC   expression of the tumour polypeptides and proteins in tumour cells; and
CC   to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC   library using well known techniques
XX
SQ   Sequence 207 BP; 81 A; 38 C; 54 G; 34 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 6;  Length 207;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        164 ATGTCAAAAAAAAAAAAAAAAAAAA 188


RESULT 35
ABN94089/c
ID   ABN94089 standard; DNA; 222 BP.
XX
AC   ABN94089;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #587 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 587; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 222 BP; 45 A; 46 C; 44 G; 87 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 6;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 8.3;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         39 ATGTCAAAAAAAAAAAAAAAAAAAA 15


RESULT 36
AAS56440/c
ID   AAS56440 standard; cDNA; 330 BP.
XX
AC   AAS56440;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human cDNA for an ovarian cancer protein #64.
XX
KW   Human; ss; ovarian cancer protein; cancer; tumour; ovarian cancer;
KW   endometrial cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200170976-A2.
XX
```

PD 27-SEP-2001.
XX
PF 20-MAR-2001; 2001WO-US009062.
XX
PR 21-MAR-2000; 2000US-0190710P.
PR 22-JUN-2000; 2000US-0213748P.
PR 19-DEC-2000; 2000US-0257276P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Pyle RA, Stolk JA;
XX
DR WPI; 2001-607531/69.
XX
PT Nucleic acids encoding 222 polypeptides associated with ovarian and
PT endometrial cancers, useful for diagnosing, preventing and treating
PT cancers.
XX
PS Claim 1; Page 137-138; 187pp; English.
XX
CC The invention relates to human polynucleotides encoding proteins
CC associated with ovarian and endometrial cancers. The polynucleotides and
CC the proteins they encode may be used in the prevention, diagnosis and
CC treatment of diseases associated with the inappropriate expression of
CC ovarian and endometrial cancer polypeptides (OECPs). For example, the
CC polynucleotide (or an expression vector comprising the polynucleotide)
CC and the OECP may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of OECPs by expressing inactive proteins or to
CC supplement the patients own production of them. Additionally, the
CC polynucleotide may be used to produce the OECPs, by inserting the nucleic
CC acids into a host cell and culturing the cell to express the protein. The
CC polynucleotide and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acids in samples, and therefore which patients may be in
CC need of restorative therapy. The OECPs may also be used as antigens in
CC the production of anti-OECP antibodies and in assays to identify
CC modulators of it's expression and activity. The anti-OECP antibodies and
CC antagonists may also be used to down regulate expression and activity.
CC The anti-OECP antibodies may also be used as diagnostic agents for
CC detecting the presence of OECP in samples (e.g. by enzyme linked
CC immunosorbant assay (ELISA)) and hence diagnose patients with cancers.
CC The present sequence is a ovarian and endometrial cancer linked cDNA of
CC the invention
XX
SQ Sequence 330 BP; 83 A; 76 C; 64 G; 103 T; 0 U; 4 Other;

Query Match 1.9%; Score 25; DB 4; Length 330;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 30 ATGTCAAAAAAAAAAAAAAAAAAAA 6

RESULT 37
ABX36291/c
ID ABX36291 standard; cDNA; 354 BP.
XX
AC ABX36291;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #1456.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 1456; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 354 BP; 92 A; 65 C; 92 G; 105 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 7; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1302 GATGTCAAAAAAAAAAAAAAAAAAA 1326
||||||||||||||||||||||||||
Db 39 GATGTCAAAAAAAAAAAAAAAAAAA 15

RESULT 38
AAL09871/c
ID AAL09871 standard; cDNA; 365 BP.
XX
AC AAL09871;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 2328.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX

OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 446; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 365 BP; 99 A; 78 C; 65 G; 111 T; 0 U; 12 Other;

Query Match 1.9%; Score 25; DB 4; Length 365;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 86 ATGTCAAAAAAAAAAAAAAAAAAAA 62

RESULT 39
ACC48717
ID ACC48717 standard; cDNA; 390 BP.
XX
AC ACC48717;
XX
DT 11-AUG-2003 (first entry)
XX
DE Mouse wound healing biomarker EST AA608387 103723_at.
XX
KW Mouse; wound healing; biomarker; marker; vulnerary; vasotropic;
KW antipsoriatic; cytostatic; ophthalmological; antiulcer; gene therapy;
KW expressed sequence tag; EST; ss.
XX
OS Mus musculus.
XX
PN EP1270044-A2.
XX
PD 02-JAN-2003.
XX
PF 17-JUN-2002; 2002EP-00254207.
XX
PR 18-JUN-2001; 2001GB-00014869.
PR 13-JUL-2001; 2001US-0305346P.
XX
PA (PFIZ) PFIZER LTD.
PA (PFIZ) PFIZER INC.
XX
PI Burslem MF, Cooper L, Johnson CM, Martin P;
XX
DR WPI; 2003-291795/29.
XX
PT Novel wound healing biomarkers including genes and protein products
PT identified as being involved in wound healing purposes, useful for
PT identifying compounds for treating wounds, wound healing disorders or
PT inflammation.
XX
PS Claim 1; Page 165; 176pp; English.
XX
CC The present DNA sequence corresponds to an expressed sequence tag (EST)
CC identified from an analysis of murine genes that are up- or down-
CC regulated during a wound healing response, with or without the
CC inflammatory response usually occurring in healing wounds. The gene that
CC encodes this EST encodes a protein important in the late, advanced phase
CC of wound healing, i.e. in the cessation of the healing response,
CC maturation and remodelling of the repaired tissue. A gene comprising this
CC EST, a protein encoded by the gene, the human homologue of the gene, and
CC a biomarker derived from the gene are claimed. Antagonists and inhibitors
CC of this novel target (e.g. antibodies, antisense oligonucleotides and
CC ribozymes) can be used to treat wounds or disorders characterised by
CC excessive wound healing, such as scarring, fibrosis, restenosis post-
CC angioplasty, post-traumatic/surgical adhesions of the peritoneal cavity,
CC joints and ligaments, psoriasis, benign prostatic hyperplasia, glaucoma
CC or peripheral nerve injury. Agonists or activators of this target (e.g.
CC the target protein delivered by means of gene therapy) can be used to
CC treat a disease or disorder characterised by impaired healing response,
CC such as chronic dermal ulcers, oral mucocystis, emphysema, ulcerative
CC diseases of the gastrointestinal tract and cystitis. Methods are provided
CC for monitoring the progress of wound healing and for identification of
CC individuals with wound healing disorders
XX
SQ Sequence 390 BP; 143 A; 59 C; 72 G; 116 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 7; Length 390;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 356 ATGTCAAAAAAAAAAAAAAAAAAAA 380

RESULT 40
AAL18543/c
ID AAL18543 standard; cDNA; 414 BP.
XX
AC AAL18543;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 11000.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 1959; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 414 BP; 117 A; 83 C; 73 G; 140 T; 0 U; 1 Other;

Query Match 1.9%; Score 25; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 47 ATGTCAAAAAAAAAAAAAAAAAAAA 23

RESULT 41
ABV18188
ID ABV18188 standard; cDNA; 425 BP.
XX
AC ABV18188;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 18179.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3000; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 425 BP; 185 A; 75 C; 40 G; 89 T; 0 U; 36 Other;

Query Match 1.9%; Score 25; DB 5; Length 425;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 131 ATGTCAAAAAAAAAAAAAAAAAAAA 155

RESULT 42
ACH20487
ID ACH20487 standard; cDNA; 430 BP.
XX
AC ACH20487;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult liver cDNA #99.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 7699; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,

```
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 430 BP; 117 A; 110 C; 100 G; 103 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 8;  Length 430;
  Best Local Similarity   100.0%;  Pred. No. 7.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        405 ATGTCAAAAAAAAAAAAAAAAAAAA 429


RESULT 43
ACF03901
ID   ACF03901 standard; cDNA; 433 BP.
XX
AC   ACF03901;
XX
DT   19-SEP-2003  (first entry)
XX
DE   Rice endosperm expressed sequence tag (EST) cDNA SEQ ID NO:6.
XX
KW   Rice; plant; endosperm; expressed sequence tag; EST; gene chip;
KW   expression sequence label; gene; ss.
XX
OS   Oryza sativa.
XX
PN   CN1366069-A.
XX
PD   28-AUG-2002.
XX
PF   31-OCT-2001; 2001CN-00135857.
XX
PR   31-OCT-2001; 2001CN-00135857.
XX
PA   (UYZH-) UNIV ZHEJIANG.
XX
PI   Dong H,  Li D;
XX
DR   WPI; 2003-382546/37.
XX
PT   Rice endosperm specific expression sequence label and gene chip prepared
PT   by it.
XX
PS   Claim 1; Page 12 (Disclosure); 28pp; Chinese.
XX
CC   The present invention describes the rice endosperm specific expressed
CC   sequence tag (EST) cDNA sequences given in ACF03896 to ACF03943 (SEQ ID
CC   NO:1 to 50, where SEQ ID NO:16 and 44 have not been given in the
CC   specification), isolated from a rice cDNA library. The ESTs are described
CC   as expression sequence labels. Also described are gene chips constituted
CC   from the expression sequence labels. The expression sequence labels can
CC   be used for DNA sequencing to a configured rice endosperm cDNA library
XX
SQ   Sequence 433 BP; 107 A; 93 C; 104 G; 129 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 7;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 7.4;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        409 ATGTCAAAAAAAAAAAAAAAAAAAA 433


RESULT 44
ABX43422/c
ID   ABX43422 standard; cDNA; 464 BP.
XX
AC   ABX43422;
XX
DT   21-FEB-2003  (first entry)
XX
DE   Bovine EST associated with lactation/muscle/fat deposition #8587.
XX
KW   Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW   muscle deposition; fat deposition; genome mapping; gene identification;
KW   gene analysis; cattle breeding.
XX
OS   Bos Taurus.
XX
PN   US2002137139-A1.
XX
PD   26-SEP-2002.
XX
PF   24-SEP-2001; 2001US-00960352.
XX
PR   12-JAN-1999;   99US-0115707P.
PR   11-JAN-2000; 2000US-00480902.
XX
PA   (BYAT/) BYATT J C.
PA   (MATH/) MATHIALAGAN N.
PA   (TAON/) TAO N.
PA   (WARR/) WARREN W C.
XX
PI   Byatt JC,  Mathialagan N,  Tao N,  Warren WC;
XX
DR   WPI; 2003-110599/10.
XX
PT   New nucleic acid associated with lactation, and muscle and fat
PT   deposition, useful for genome mapping, gene identification and analysis,
PT   cattle breeding, or for genetically improving cattle.
XX
PS   Claim 2; SEQ ID NO 8587; 245pp; English.
XX
CC   The invention relates to a purified nucleic acid molecule associated with
CC   lactation or muscle and fat deposition (designated LMFD), derived from
CC   cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC   nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC   appearing as ABX34836-ABX49947, or complements of them. Also included are
CC   ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC   acid linked to a promoter and a 3' non- translated sequence that
CC   functions in the cell to cause termination of transcription and addition
CC   of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC   (2) determining a level or pattern of a molecule in a bovine cell or
CC   tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC   of the 15112 nucleic acid sequences or its complement or fragment) with a
CC   complementary nucleic acid molecule obtained from the bovine cell or
CC   tissue, where hybridisation between the marker nucleic acid and the
CC   complementary nucleic acid permits the detection of the molecule; and (b)
CC   detecting the level or pattern of the complementary nucleic acid, where
CC   the detection of the complementary nucleic acid is predictive of the
CC   level or pattern of the molecule. The LMFD nucleic acid is used for
CC   determining a level or pattern of a molecule in a bovine cell or tissue.
CC   It is useful for genome mapping, gene identification and analysis, cattle
CC   breeding, preparation of constructs for use in cattle gene expression, or
CC   for genetically improving cattle. The present sequence is one of the
CC   15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC   present sequence was not shown in the specification but was obtained in
CC   electronic format from the USPTO web site:
CC   seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ   Sequence 464 BP; 132 A; 85 C; 107 G; 140 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 7;  Length 464;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAA 1326
```

```
                    ||||||||||||||||||||||
Db          42 GATGTCAAAAAAAAAAAAAAAAAAA 18
```

RESULT 45
AAQ74392
ID AAQ74392 standard; DNA; 469 BP.
XX
AC AAQ74392;
XX
DT 25-MAR-2003 (revised)
DT 03-JUN-1995 (first entry)
XX
DE Clone pSP65-XaI-4 encoding leech factor Xa inhibitor.
XX
KW Leech; factor Xa; inhibitor; ss.
XX
OS Hirudo medicinalis.
XX
FH Key   Location/Qualifiers
FT CDS   10. .469
FT   /*tag= a
XX
PN WO9423735-A1.
XX
PD 27-OCT-1994.
XX
PF 08-APR-1994; 94WO-US003871.
XX
PR 09-APR-1993; 93US-00045805.
XX
PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
PI Zeelon EP, Werber MM, Levanon A;
XX
DR WPI; 1994-341477/42.
DR P-PSDB; AAR62516.
XX
PT Polypeptide from Hirudo medicinalis - has factor Xa inhibitory activity,
PT useful for treating blood coagulation disorders.
XX
PS Example; Fig 5; 135pp; English.
XX
CC Total RNA was extracted from 120 leeches, Hirudo medicinalis. An aliquot
CC of the poly A+ mRNA was used as template in a reverse transcription in
CC the presence of the synthetic primer A (AAQ74385) which provides
CC complementarity to the polyA sequence of the various mRNAs. Following
CC synthesis of the ss cDNA, the mRNA was degraded. An aliquot of the
CC neutralised ss cDNA was then subjected to PCR amplification using as
CC reverse primer the synthetic degenerative DNA oligomer B (AAQ74394). This
CC synthetic primer was prepd. in accordance with the first nine N-terminal
CC AAs of the naturally occurring FXa inhibitor and encodes the first nine N
CC -terminal AAs in AAR62518. The PCR amplification products were loaded
CC onto agarose gel. Three distinct bands of about 350bp, 450 bp and 700 bp
CC were observed. The bands were hybridised to a synthetic radiolabeled DNA
CC probe (probe C) corresp. to N-terminal AAs 14-19 of AAR62518. The three
CC PCR products were found to hybridise with probe C. However, the band
CC corresp. to 700 bp was found to hybridise relatively poorly to the 350 bp
CC and 450 bp fragments. The DNA was purified, digested, and subcloned into
CC pSP65 which was used to transform E.coli strain MC1061. Transformants
CC were screening using probe C. Plasmid DNA was prepd. from positive
CC clones. Two classes of clones were obtd.: (a) those with an insert of
CC about 290 bps; and (b) those with an insert of about 450 bps. pSP65-XaI-4
CC contains an insert of about 450 bps. Its DNA sequence and deduced AA
CC sequence are given in AAQ74392/R62516. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

  Query Match   1.9%; Score 25; DB 2; Length 469;
  Best Local Similarity 100.0%; Pred. No. 7.3;
  Matches   25; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAA 1327
                ||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAA 469
```

RESULT 46
AAQ72957
ID AAQ72957 standard; DNA; 469 BP.
XX
AC AAQ72957;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-JUN-1995 (first entry)
XX
DE Leech PCR clone pSP65-XaI-4 encoding factor Xa inhibitor (FXaI).
XX
KW Leech; factor Xa inhibitor; ss.
XX
OS Hirudo medicinalis.
XX
FH Key   Location/Qualifiers
FT CDS   10. .468
FT   /*tag= a
XX
PN WO9423709-A1.
XX
PD 27-OCT-1994.
XX
PF 08-APR-1994; 94WO-US003918.
XX
PR 09-APR-1993; 93US-00045804.
XX
PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
PA (YISS) YISSUM RES & DEV CO.
XX
PI Werber MM, Zeelon EP, Levanon A, Guy R, Goldlust A, Rigbi M;
PI Panet A, Fischer M;
XX
DR WPI; 1994-341457/42.
DR P-PSDB; AAR62624.
XX
PT Recombinant factor 10a inhibitor of Hirudo medicinalis - for treating
PT excessive blood coagulation, partic. thrombosis, also related DNA,
PT vectors, transformed cells and antibodies.
XX
PS Example; Fig 8; 107pp; English.
XX
CC Total RNA was extracted from 120 leeches, poly A+ mRNA was isolated and
CC used as template in a reverse transcription reaction in the presence of
CC primer A (AAQ72958), which provides complementarity to the poly A
CC sequence of the various mRNAs. Following synthesis of the ss cDNA, the
CC mRNA was degraded and the ss cDNA subjected to PCR amplification using as
CC reverse primer the degenerative oligo B (AAQ72959) which was prepd. in
CC accordance with the first nine N- terminal AAs of Factor Xa inhibitor
CC (FXaI) extracted from leech saliva. The PCR amplification products were
CC loaded onto agarose gel, the bands blotted onto nitrocellulose paper and
CC hybridised to a synthetic radiolabeled DNA probe (probe C (AAQ72962))
CC corresp. to N- terminal AAs 14-19 of leech saliva FXaI. Probe C was also
CC used to screen transformants of the PCR amplified DNA. The DNA sequence
CC and deduced AA sequence of one of the positive clones pSP65-XaI-4 (clone
CC 4) is shown in AAQ72957/R62624. It is noted that met 5 in AAR62624 is the
CC fourth AA (Met 1 is added by the bacterial host cell) of the FSaI
CC polypeptide which differs from Ile4 which is the fourth AA of the N-
CC terminal sequence of the naturally occurring FXaI isolated from leech
CC saliva. This discrepancy is apparently due to a mistake by the DNA
CC polymerase in the course of the PCR. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

  Query Match   1.9%; Score 25; DB 2; Length 469;
  Best Local Similarity 100.0%; Pred. No. 7.3;

```
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 47
AAV33944
ID   AAV33944 standard; DNA; 469 BP.
XX
AC   AAV33944;
XX
DT   25-JAN-1999  (first entry)
XX
DE   Leech FXaI gene clone pSP65-XaI-4.
XX
KW   European leech; Factor Xa inhibitor; FXaI; probe; hybridisation; RT-PCR;
KW   influenza virus; infection; coagulation; primer; amplification; ss.
XX
OS   Hirudo medicinalis.
XX
FH   Key             Location/Qualifiers
FT   CDS             10. .468
FT                   /*tag=  a
FT                   /transl_except= (pos:268. .270, aa:Xaa)
FT                   /transl_except= (pos:280. .282, aa:Xaa)
FT                   /transl_except= (pos:328. .330, aa:Xaa)
FT                   /transl_except= (pos:334. .336, aa:Xaa)
FT                   /transl_except= (pos:391. .393, aa:Xaa)
FT                   /note= "Xaa is unknown and not given in the corresponding
FT                   protein sequence"
XX
PN   US5824641-A.
XX
PD   20-OCT-1998.
XX
PF   06-JUN-1995;   95US-00467389.
XX
PR   09-APR-1993;   93US-00045805.
PR   08-APR-1994;   94US-00225442.
XX
PA   (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
PI   Werber MM;
XX
DR   WPI; 1998-582592/49.
DR   P-PSDB; AAW68544.
XX
PT   Treating or preventing influenza - comprises the administration of factor
PT   Xa inhibiting polypeptide.
XX
PS   Example 3; Fig 5; 56pp; English.
XX
CC   This sequence represents clone pSP65-XaI-4 which was amplified from DNA
CC   isolated from the European leech (Hirudo medicinalis). The open reading
CC   frame contains 5 internal stop codons when decoded from the first ATG
CC   initiation codon. No equivalent residues are given in the corresponding
CC   protein sequence. The fragment is a primary isolated and partially
CC   encoded the leech Factor Xa inhibitor (FXaI). The fragment was amplified
CC   by the primers AAV33941-V33942. The fragment was used as a probe to
CC   screen a cDNA library for the FXaI gene (AAV33940). The FXaI protein is
CC   claimed to be useful in the treatment of influenza viral infection by
CC   preventing the infection and re-infection cycle of cells by the virus. An
CC   activation enzyme implicated in the viral infection cycle has been shown
CC   to be similar to the chicken coagulation Factor Xa and it is thought that
CC   the corresponding human protein may act in a similar manner. Thus
CC   inhibitors of Factor Xa may prevent influenza infections
XX
SQ   Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

 Query Match             1.9%;  Score 25;  DB 2;  Length 469;
 Best Local Similarity   100.0%;  Pred. No. 7.3;
```

```
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 48
AAX01821
ID   AAX01821 standard; DNA; 469 BP.
XX
AC   AAX01821;
XX
DT   13-APR-1999  (first entry)
XX
DE   H. medicinalis FXaI clone pSP65-XaI-4 DNA.
XX
KW   Factor XA; FXaI; inhibitor; blood; coagulation; European leech;
KW   post-operative trauma; obesity; pregnancy; oral contraceptive; stroke;
KW   cerebrovascular disorder; ss.
XX
OS   Hirudo medicinalis.
XX
FH   Key             Location/Qualifiers
FT   CDS             10. .469
FT                   /*tag=  a
FT                   /product= "pSP65-XaI-4"
FT                   /note= "contains internal stop codons"
XX
PN   US5863534-A.
XX
PD   26-JAN-1999.
XX
PF   06-JUN-1995;   95US-00469219.
XX
PR   09-APR-1993;   93US-00045805.
PR   08-APR-1994;   94US-00225442.
XX
PA   (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
PI   Werber MM,  Levanon A,  Zeelon EP;
XX
DR   WPI; 1999-131254/11.
DR   P-PSDB; AAW92370.
XX
PT   Reducing blood coagulation - using factor Xa inhibitor polypeptides
PT   obtained from saliva of the European leech Hirudo medicinalis.
XX
PS   Disclosure; Fig 5; 56pp; English.
XX
CC   This sequence encodes a novel FXaI protein, isolated from the European
CC   leech (Hirudo medicinalis) clone pSP65-XaI-4 which inhibits Factor XA
CC   activity. This fragment is used in the construction of a polypeptide
CC   which can be used to reduce the extent of blood coagulation and has the
CC   sequence X-Y-Cys Gln Glu Glu Glu Cys Pro Asp Pro Tyr Leu Cys Ser Pro Val
CC   Thr Asn Arg Cys Glu Cys Thr Pro Val Leu Cys Arg Met Tyr Cys Lys Phe Trp
CC   Ala Lys Cys Glu Glu Leu Cys Gln Ser Ser Val Thr Asn Arg Cys Asp Cys Gln
CC   Asp Phe Lys Cys Pro Gln Ser Tyr Cys-Z, where X is Met or absent; Y is 0-
CC   29 amino acids of sequence Lys Met Cys Trp Asn Lys Gly Cys Pro Cys Gly
CC   Gln Arg Cys Asn Leu His Arg Asn Glu Cys Glu Val Ile Ala Glu Asn Ile Glu
CC   with the proviso that if Y is > 0 amino acids, then the carboxy-terminal
CC   sequence of Y beginning with Glu as the carboxy-terminal amino acid is
CC   present and where Val24 may be preceded by Gly, Z is absent or all or a
CC   part of the sequence Pro110-Lys156 (Pro His Gly Phe Glu Thr Asp Glu Asn
CC   Glu Cys Glu Val Cys Ile Cys Lys Lys Pro Thr Cys Ala Asn Cys Gly Lys Thr
CC   Thr Lys Lys Lys Pro Arg Thr Ile Asp Arg Leu Lys Asn Trp Phe Lys Lys Lys
CC   Phe Gly Lys), with the proviso that if Z is > 0 amino acid, then the
CC   amino-terminal sequence of Z beginning with Pro as the amino-terminal
CC   amino acid is present. The method can be used for treating subjects with
CC   excessive blood coagulation disorders, e.g. vascular disorders, post-
CC   operative trauma, obesity, pregnancy, side affects of oral
CC   contraceptives, prolonged immobilisation, or a cerebrovascular disorder
CC   e.g. stroke
```

```
XX
SQ   Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 49
AAV55665
ID   AAV55665 standard; DNA; 469 BP.
XX
AC   AAV55665;
XX
DT   18-MAR-1999  (first entry)
XX
DE   Factor Xa inhibitor peptide coding sequence.
XX
KW   Factor Xa inhibitor; leech; blood coagulation; thrombosis;
KW   recurrent influenza infection; ss.
XX
OS   Hirudo medicinalis.
XX
PN   US5858970-A.
XX
PD   12-JAN-1999.
XX
PF   07-JAN-1997;   97US-00779379.
XX
PR   09-APR-1993;   93US-00045805.
PR   08-APR-1994;   94US-00225442.
XX
PA   (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
PI   Levanon A,  Zeelon EP,  Werber MM;
XX
DR   WPI; 1999-119954/10.
DR   P-PSDB; AAW73598.
XX
PT   Factor Xa inhibitory polypeptide - derived from Hirudo medicinalis,
PT   useful as anticoagulant or antithrombotic agent.
XX
PS   Example 3; Fig 5; 57pp; English.
XX
CC   This sequence encodes the full length factor Xa inhibitor peptide of the
CC   invention. The factor Xa sequence was isolated from the leech Hirudo
CC   medicinalis. It can be used to reduce the extent of blood coagulation or
CC   thrombosis or to prevent recurrent influenza infection
XX
SQ   Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 25;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 50
AAD04138
ID   AAD04138 standard; DNA; 469 BP.
XX
AC   AAD04138;
XX
DT   02-JUL-2001  (first entry)
XX
DE   DNA clone pSP65-XaI-4 encoding leech FXa inhibitor.
XX
KW   FXaI; blood coagulation; factor FXa; inhibitor; therapy; thrombosis;
KW   influenza; leech; ds.
XX
OS   Hirudo medicinalis.
XX
FH   Key             Location/Qualifiers
FT   CDS             10. .453
FT                   /*tag=  a
FT                   /product= "Leech FXa inhibitor encoded by clone pSP65-XaI
FT                   -4"
FT                   /transl_except= (pos:265. .273, aa:Gly-Thr)
FT                   /note= "Insertion of an in-frame stop codon alters the
FT                   reading frame"
FT                   /transl_except= (pos:277. .285, aa:Ile-Lys)
FT                   /note= "Insertion of an in-frame stop codon alters the
FT                   reading frame"
FT                   /transl_except= (pos:325. .333, aa:Leu-Arg)
FT                   /note= "Insertion of an in-frame stop codon alters the
FT                   reading frame"
FT                   /transl_except= (pos:331. .339, aa:Arg-Ile)
FT                   /note= "Insertion of an in-frame stop codon alters the
FT                   reading frame"
FT                   /transl_except= (pos:388. .396, aa:Ser-Ile)
FT                   /note= "Insertion of an in-frame stop codon alters the
FT                   reading frame; All the above translation exceptions
FT                   correspond to the sequence shown in AAE00818"
FT                   /transl_except= (pos:268. .270, aa:Xaa)
FT                   /transl_except= (pos:280. .282, aa:Xaa)
FT                   /transl_except= (pos:328. .330, aa:Xaa)
FT                   /transl_except= (pos:334. .336, aa:Xaa)
FT                   /transl_except= (pos:391. .393, aa:Xaa)
FT                   /note= "Xaa corresponds to in-frame stop codon; the above
FT                   translation exceptions correspond to the sequence shown
FT                   in AAE00755; CDS does not include stop codon"
FT                   /partial
XX
PN   US6211341-B1.
XX
PD   03-APR-2001.
XX
PF   11-JAN-1999;   99US-00228152.
XX
PR   09-APR-1993;   93US-00045805.
PR   08-APR-1994;   94US-00225442.
PR   06-JUN-1995;   95US-00469219.
XX
PA   (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX
PI   Zeelon EP,  Werber MM,  Levanon A;
XX
DR   WPI; 2001-280869/29.
DR   P-PSDB; AAE00755, AAE00818.
XX
PT   New antibody reactive with an epitope of a polypeptide having factor Xa
PT   inhibitory activity, useful for purifying or detecting a polypeptide
PT   having Factor Xa inhibitory activity.
XX
PS   Example 3; Col 51-54; 57pp; English.
XX
CC   The invention relates to an antibody which specifically reacts with an
CC   epitope of a novel (unknown) polypeptide derived from Hirudo medicinalis.
CC   This polypeptide is an inhibitor of blood coagulation factor FXa and is
CC   designated as FXaI. The polypeptide of the invention is used in
CC   diagnostic and therapeutic methods applied in conditions characterised by
CC   the excessive blood coagulation and thrombosis. The FXaI polypeptide is
CC   also used to prevent recurrent influenza infection. The monoclonal
CC   antibody specific for FXaI polypeptide is useful for purifying or
CC   detecting FXaI. The present sequence is DNA clone pSP65-XaI-4 encoding
CC   leech FXa inhibitor
XX
SQ   Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;
```

Query Match 1.9%; Score 25; DB 4; Length 469;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469
```

Search completed: September 17, 2004, 03:18:25
Job time : 564 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 03:09:09 ; Search time 3727 Seconds
(without alignments)
1797.146 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtggtcacaggtag.....aaaaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3327077 seqs, 2523723180 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1327	100.0	3826	9	US-09-927-091-3	Sequence 3, Appli
	2	1308	98.6	23433	9	US-09-927-091-7	Sequence 7, Appli
	3	1257	94.7	30676	9	US-09-927-091-8	Sequence 8, Appli
	4	726	54.7	30625	9	US-09-927-091-5	Sequence 5, Appli
C	5	420	31.7	610	13	US-10-027-632-100265	Sequence 100265,
C	6	420	31.7	610	16	US-10-027-632-100265	Sequence 100265,
	7	60	4.5	60	10	US-09-908-975-16197	Sequence 16197, A
C	8	29	2.2	1652	16	US-10-116-275-299	Sequence 299, App
	9	27	2.0	1754	17	US-10-437-963-80513	Sequence 80513, A
C	10	27	2.0	248436	13	US-10-087-192-2014	Sequence 2014, Ap
	11	26	2.0	273	12	US-09-969-034-4195	Sequence 4195, Ap
C	12	26	2.0	393	9	US-09-960-352-6255	Sequence 6255, Ap
C	13	26	2.0	411	9	US-09-960-352-13315	Sequence 13315, A
	14	26	2.0	4702	16	US-10-295-027-141	Sequence 141, App
	15	26	2.0	4702	16	US-10-295-027-1164	Sequence 1164, Ap
	16	26	2.0	6808	17	US-10-437-963-87284	Sequence 87284, A
	17	25	1.9	136	13	US-10-085-783A-47925	Sequence 47925, A
	18	25	1.9	136	16	US-10-242-535A-47925	Sequence 47925, A
	19	25	1.9	143	13	US-10-085-783A-51657	Sequence 51657, A
	20	25	1.9	143	16	US-10-242-535A-51657	Sequence 51657, A
	21	25	1.9	156	13	US-10-085-783A-45661	Sequence 45661, A
	22	25	1.9	156	16	US-10-242-535A-45661	Sequence 45661, A
	23	25	1.9	184	9	US-09-867-701-9253	Sequence 9253, Ap
C	24	25	1.9	191	9	US-09-962-436-134	Sequence 134, App
C	25	25	1.9	191	9	US-09-954-456-1020	Sequence 1020, Ap
C	26	25	1.9	191	9	US-09-880-107-525	Sequence 525, App
	27	25	1.9	207	9	US-09-867-701-10779	Sequence 10779, A
C	28	25	1.9	222	9	US-09-880-107-587	Sequence 587, App
C	29	25	1.9	256	13	US-10-424-599-31507	Sequence 31507, A
	30	25	1.9	296	13	US-10-085-783A-52889	Sequence 52889, A
	31	25	1.9	296	16	US-10-242-535A-52889	Sequence 52889, A
C	32	25	1.9	330	9	US-09-813-358-64	Sequence 64, Appl
C	33	25	1.9	330	10	US-09-997-279-64	Sequence 64, Appl
C	34	25	1.9	330	12	US-09-922-293-42	Sequence 42, Appl
C	35	25	1.9	354	9	US-09-960-352-1456	Sequence 1456, Ap
	36	25	1.9	358	13	US-10-424-599-52765	Sequence 52765, A
	37	25	1.9	382	17	US-10-437-963-88420	Sequence 88420, A
	38	25	1.9	390	13	US-10-175-184A-31	Sequence 31, Appl
	39	25	1.9	430	10	US-09-918-995-7699	Sequence 7699, Ap
C	40	25	1.9	464	9	US-09-960-352-8587	Sequence 8587, Ap
C	41	25	1.9	467	15	US-10-198-846-1221	Sequence 1221, Ap
C	42	25	1.9	483	15	US-10-102-524-36	Sequence 36, Appl
	43	25	1.9	483	17	US-10-437-963-85418	Sequence 85418, A
	44	25	1.9	486	13	US-10-027-632-95705	Sequence 95705, A
	45	25	1.9	486	13	US-10-027-632-305902	Sequence 305902,
	46	25	1.9	486	16	US-10-027-632-95705	Sequence 95705, A
	47	25	1.9	486	16	US-10-027-632-305902	Sequence 305902,
C	48	25	1.9	492	17	US-10-437-963-19508	Sequence 19508, A
C	49	25	1.9	503	15	US-10-029-386-1201	Sequence 1201, Ap
	50	25	1.9	527	9	US-09-925-301-787	Sequence 787, App
	51	25	1.9	549	12	US-09-969-034-4262	Sequence 4262, Ap
C	52	25	1.9	558	17	US-10-437-963-79631	Sequence 79631, A
	53	25	1.9	591	10	US-09-764-891-1170	Sequence 1170, Ap
	54	25	1.9	593	13	US-10-027-632-96478	Sequence 96478, A
	55	25	1.9	593	16	US-10-027-632-96478	Sequence 96478, A
	56	25	1.9	613	17	US-10-437-963-17992	Sequence 17992, A
	57	25	1.9	616	13	US-10-027-632-308292	Sequence 308292,
	58	25	1.9	616	13	US-10-027-632-308293	Sequence 308293,
	59	25	1.9	616	13	US-10-027-632-308294	Sequence 308294,
	60	25	1.9	616	13	US-10-027-632-308295	Sequence 308295,
	61	25	1.9	616	16	US-10-027-632-308292	Sequence 308292,
	62	25	1.9	616	16	US-10-027-632-308293	Sequence 308293,
	63	25	1.9	616	16	US-10-027-632-308294	Sequence 308294,
	64	25	1.9	616	16	US-10-027-632-308295	Sequence 308295,
C	65	25	1.9	617	15	US-10-198-846-8111	Sequence 8111, Ap
	66	25	1.9	626	13	US-10-027-632-192324	Sequence 192324,
	67	25	1.9	626	16	US-10-027-632-192324	Sequence 192324,
	68	25	1.9	645	13	US-09-823-245A-41	Sequence 41, Appl
C	69	25	1.9	648	15	US-10-198-846-8765	Sequence 8765, Ap
C	70	25	1.9	651	17	US-10-437-963-92674	Sequence 92674, A
C	71	25	1.9	708	17	US-10-767-701-11186	Sequence 11186, A
C	72	25	1.9	730	13	US-10-027-632-100233	Sequence 100233,
C	73	25	1.9	730	16	US-10-027-632-100233	Sequence 100233,
	74	25	1.9	732	17	US-10-437-963-46535	Sequence 46535, A
	75	25	1.9	749	13	US-10-027-632-20596	Sequence 20596, A
	76	25	1.9	749	16	US-10-027-632-20596	Sequence 20596, A
	77	25	1.9	834	13	US-10-027-632-165337	Sequence 165337,
	78	25	1.9	834	13	US-10-027-632-165338	Sequence 165338,
	79	25	1.9	834	16	US-10-027-632-165337	Sequence 165337,
	80	25	1.9	834	16	US-10-027-632-165338	Sequence 165338,
	81	25	1.9	1026	17	US-10-437-963-50383	Sequence 50383, A
	82	25	1.9	1092	9	US-09-925-301-317	Sequence 317, App
	83	25	1.9	1163	17	US-10-437-963-9930	Sequence 9930, Ap
	84	25	1.9	1193	13	US-10-424-599-75445	Sequence 75445, A
	85	25	1.9	1307	15	US-10-198-846-10951	Sequence 10951, A
C	86	25	1.9	1410	17	US-10-437-963-10738	Sequence 10738, A
	87	25	1.9	1451	15	US-10-106-698-401	Sequence 401, App

	88	25	1.9	1452	9	US-09-925-299-183 Sequence 183, App
	89	25	1.9	1452	10	US-09-925-299-183 Sequence 183, App
	90	25	1.9	1624	9	US-09-764-864-126 Sequence 126, App
	91	25	1.9	1696	9	US-09-925-300-649 Sequence 649, App
	92	25	1.9	1781	13	US-10-424-599-24574 Sequence 24574, A
c	93	25	1.9	1790	9	US-09-867-550-1519 Sequence 1519, Ap
	94	25	1.9	1859	17	US-10-615-659-10 Sequence 10, Appl
	95	25	1.9	1859	17	US-10-635-977-10 Sequence 10, Appl
	96	25	1.9	2283	15	US-10-198-846-10281 Sequence 10281, A
	97	25	1.9	2484	17	US-10-437-963-2157 Sequence 2157, Ap
c	98	25	1.9	2511	17	US-10-437-963-100661 Sequence 100661,
	99	25	1.9	2522	9	US-09-745-763-39 Sequence 39, Appl
c	100	25	1.9	2586	16	US-10-104-047-875 Sequence 875, App
	101	25	1.9	2629	15	US-10-106-698-1916 Sequence 1916, Ap
	102	25	1.9	2796	9	US-09-529-063-51 Sequence 51, Appl
	103	25	1.9	2796	15	US-10-414-378-51 Sequence 51, Appl
	104	25	1.9	2806	13	US-10-424-599-137631 Sequence 137631,
	105	25	1.9	3465	17	US-10-615-659-11 Sequence 11, Appl
	106	25	1.9	3465	17	US-10-635-977-11 Sequence 11, Appl
	107	25	1.9	3504	17	US-10-437-963-54840 Sequence 54840, A
	108	25	1.9	3554	17	US-10-615-659-12 Sequence 12, Appl
	109	25	1.9	3554	17	US-10-635-977-12 Sequence 12, Appl
	110	25	1.9	4572	17	US-10-450-826-108 Sequence 108, App
	111	25	1.9	4740	9	US-09-945-265-3 Sequence 3, Appli
	112	25	1.9	4740	10	US-09-902-481A-2 Sequence 2, Appli
	113	25	1.9	4740	15	US-10-207-655-175 Sequence 175, App
	114	25	1.9	4740	15	US-10-136-819-18 Sequence 18, Appl
	115	25	1.9	4740	16	US-10-159-563-339 Sequence 339, App
	116	25	1.9	5307	13	US-10-342-887-370 Sequence 370, App
	117	25	1.9	5307	13	US-10-172-118-370 Sequence 370, App
	118	25	1.9	5307	14	US-10-028-056-2 Sequence 2, Appli
	119	25	1.9	5307	17	US-10-429-160-37 Sequence 37, Appl
	120	25	1.9	5808	16	US-10-295-027-139 Sequence 139, App
	121	25	1.9	5808	16	US-10-295-027-984 Sequence 984, App
c	122	25	1.9	23611	17	US-10-322-281-61 Sequence 61, Appl
c	123	25	1.9	35548	13	US-10-087-192-1786 Sequence 1786, Ap
c	124	25	1.9	57501	17	US-10-304-126-11 Sequence 11, Appl
c	125	25	1.9	97247	13	US-10-087-192-688 Sequence 688, App
c	126	25	1.9	126001	16	US-10-175-492-13 Sequence 13, Appl
	127	25	1.9	130877	17	US-10-322-281-54 Sequence 54, Appl
c	128	25	1.9	196063	17	US-10-322-281-612 Sequence 612, App
	129	25	1.9	312477	17	US-10-317-883A-12 Sequence 12, Appl
c	130	25	1.9	357652	17	US-10-322-696-34 Sequence 34, Appl
	131	24	1.8	43	12	US-10-461-126-4 Sequence 4, Appli
c	132	24	1.8	92	15	US-10-060-036-1858 Sequence 1858, Ap
	133	24	1.8	158	9	US-09-867-701-9919 Sequence 9919, Ap
	134	24	1.8	173	9	US-09-960-352-11302 Sequence 11302, A
c	135	24	1.8	201	17	US-10-741-601-6995 Sequence 6995, Ap
	136	24	1.8	203	13	US-10-424-599-46838 Sequence 46838, A
c	137	24	1.8	206	9	US-09-969-708-588 Sequence 588, App
c	138	24	1.8	206	9	US-09-880-107-2731 Sequence 2731, Ap
c	139	24	1.8	229	9	US-09-983-965-4546 Sequence 4546, Ap
	140	24	1.8	251	13	US-10-424-599-46211 Sequence 46211, A
	141	24	1.8	289	9	US-09-777-564-153 Sequence 153, App
	142	24	1.8	289	15	US-10-015-219-153 Sequence 153, App
c	143	24	1.8	299	9	US-09-814-122-28 Sequence 28, Appl
c	144	24	1.8	299	13	US-10-649-857-28 Sequence 28, Appl
	145	24	1.8	319	13	US-10-424-599-20333 Sequence 20333, A
	146	24	1.8	326	13	US-10-424-599-44379 Sequence 44379, A
	147	24	1.8	349	13	US-10-424-599-103537 Sequence 103537,
	148	24	1.8	357	13	US-10-424-599-4025 Sequence 4025, Ap
	149	24	1.8	359	13	US-10-085-783A-53994 Sequence 53994, A
	150	24	1.8	359	16	US-10-242-535A-53994 Sequence 53994, A

ALIGNMENTS

RESULT 1
US-09-927-091-3
; Sequence 3, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTT, STEVE
; APPLICANT: CHANDLER, DAWN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3826
; TYPE: DNA
; ORGANISM: Human
US-09-927-091-3

Query Match 100.0%; Score 1327; DB 9; Length 3826;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
Db    2500 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 2559

Qy      61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
Db    2560 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 2619

Qy     121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
Db    2620 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 2679

Qy     181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
Db    2680 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 2739

Qy     241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
Db    2740 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2799

Qy     301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
Db    2800 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 2859

Qy     361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
Db    2860 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 2919

Qy     421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
Db    2920 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 2979

Qy     481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
Db    2980 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 3039

Qy     541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
Db    3040 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 3099

Qy     601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
Db    3100 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 3159

Qy     661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
Db    3160 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 3219

Qy     721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
```

```
Db       3220 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 3279

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3280 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 3339

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3340 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 3399

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3400 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 3459

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3460 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 3519

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3520 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 3579

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3580 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 3639

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3640 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 3699

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3700 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 3759

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3760 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCAAAAAAAAAAAAA 3819

Qy       1321 AAAAAAA 1327
              |||||||
Db       3820 AAAAAAA 3826
```

```
RESULT 2
US-09-927-091-7
; Sequence 7, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
;  APPLICANT: KILLARY, ANN
;  APPLICANT:  LOTT, STEVE
;  APPLICANT:  CHANDLER, DAWN
;  TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
;  FILE REFERENCE: UTSC:651US
;  CURRENT APPLICATION NUMBER: US/09/927,091
;  CURRENT FILING DATE:  2001-08-09
;  PRIOR APPLICATION NUMBER: 60/227,560
;  PRIOR FILING DATE: 2000-08-23
;  PRIOR APPLICATION NUMBER: 60/225,033
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 23433
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (5071)..(23433)
;   OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-7

  Query Match           98.6%;  Score 1308;  DB 9;  Length 23433;
  Best Local Similarity  100.0%;  Pred. No. 0;
```

```
  Matches 1308;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13550 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 13609

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13610 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 13669

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13670 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 13729

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13730 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 13789

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13790 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 13849

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13850 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 13909

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13910 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 13969

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13970 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 14029

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14030 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 14089

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14090 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 14149

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14150 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 14209

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14210 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 14269

Qy        721 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14270 CTGAGGCCTGCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 14329

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14330 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 14389

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14390 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 14449

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14450 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 14509

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14510 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 14569

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14570 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 14629
```

```
Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14630 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 14689

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14690 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 14749

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14750 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 14809

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      14810 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 14857
```

RESULT 3
US-09-927-091-8
; Sequence 8, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTT, STEVE
; APPLICANT: CHANDLER, DAWN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 30676
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (6671)..(30676)
; OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-8

Query Match 94.7%; Score 1257; DB 9; Length 30676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24440 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 24499

Qy         61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24500 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 24559

Qy        121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24560 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 24619

Qy        181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24620 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 24679

Qy        241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24680 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 24739

Qy        301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24740 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 24799

Qy        361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24800 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 24859

Qy        421 AGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24860 GGACCCCTTGTCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTGCAGAAAACAAGAGC 24919

Qy        481 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24920 CACTTGTAGCTGGTTTAATTAGACAAGGATTTACTACCTGGCCCCTGGTGGCTTGCAAAA 24979

Qy        541 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24980 TTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGGAACTCCCAGCGCCAGAT 25039

Qy        601 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25040 TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA 25099

Qy        661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25100 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAATAGATGTC 25159

Qy        721 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25160 CTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCCCAAATCTAAATTTTTACAAGAGATT 25219

Qy        781 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25220 CTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 25279

Qy        841 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25280 CATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTGCGTTTTTCTGCCAC 25339

Qy        901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25340 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCCTGCCC 25399

Qy        961 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25400 CTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGGAGTGGGCTGCAAGACTGAGCCTA 25459

Qy       1021 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25460 AATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGG 25519

Qy       1081 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25520 TCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 25579

Qy       1141 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25580 AATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA 25639

Qy       1201 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25640 ACCCTTCAACTACCAGAATCTGGGCCACCCCAGCAGTATTTTTATTTAAAATGTTGCCCA 25699

Qy       1261 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      25700 TTTTATGAGTTATGATCAATTTGTATTAAATTAAAGTTACAGATGTCA 25747
```

RESULT 4
US-09-927-091-5
; Sequence 5, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTT, STEVE

```
;  APPLICANT:  CHANDLER, DAWN
;  TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
;  FILE REFERENCE: UTSC:651US
;  CURRENT APPLICATION NUMBER: US/09/927,091
;  CURRENT FILING DATE:  2001-08-09
;  PRIOR APPLICATION NUMBER: 60/227,560
;  PRIOR FILING DATE: 2000-08-23
;  PRIOR APPLICATION NUMBER: 60/225,033
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 30625
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (4754)..(30625)
;   OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-5

  Query Match           54.7%;  Score 726;  DB 9;  Length 30625;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 1146;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        157 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23131 TCCACACATTATAGTCATGTGCCACCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAG 23190

Qy        217 GGTATACCCAAAGCTGATGCAGAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 276
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      23191 GGTATACCCAAAGCTGATGCAAAGCCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCC 23250

Qy        277 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23251 CTGGATGATCGAGGTCCCCAGTAGCTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAG 23310

Qy        337 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23311 GCCTCTGTGACCTGCTAGGGTGCAGGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCC 23370

Qy        397 AAGCACT-GGGAGGGGCTGTTGGCTAGACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 455
              ||||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      23371 AAGCACTGGGGAGGGGCTGTTGGCTAAACCCCTTGTCAGACTTGGCATCTATCTCAGTTA 23430

Qy        456 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23431 GGATCCTGCTGCAGAAAACAAGAGCCACTTGTAGCTGGTTTAATTAGACAAGGATTTACT 23490

Qy        516 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23491 ACCTGGCCCCTGGTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGAA 23550

Qy        576 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23551 TTTCCAGGAACTCCCAGCGCCAGATTCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCC 23610

Qy        636 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23611 ATCTGCAGGAAGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCA 23670

Qy        696 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCC 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23671 CGGTCCGTGCCAGCCAATAGATGTCCTGAGGCCTGCCCCCTCTCCCACTTCACTCAGTTCC 23730

Qy        756 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23731 CAAATCTAAATTTTTACAAGAGATTCTGTTTGGGGGAACTTAAGTCAGATCCAGAACCTT 23790

Qy        816 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23791 GGCTGCAAGGGAGTCTGGGAAATGTCATTTCCCTAGAAGGAAGTTAGGGTGGGTGGAGCA 23850

Qy        876 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGA 935
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      23851 AGCCCCACCTGCGTTTTTCTGCCACAGCATCCAATCGTGAAAAACTCGGGAGAGGGTGGA 23910

Qy        936 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23911 GTCCACATCTAGGGTTGTCCTGCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGA 23970

Qy        996 GGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 1055
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      23971 GGAGTGGGCTGCAAAACTGAGCCTAAATGTCTCCCCGGCCTTGACTTTTCTTTCTAGTCC 24030

Qy       1056 TGGGGCCTAGATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 1115
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24031 TGGGGCCTAAATTCTGCACTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGA 24090

Qy       1116 AGAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 1175
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24091 AAAGCTAAACACAGGGGGTTCTTAAAATGGCTGCCCCCGCCACCCGGGCCTCCCTTGGGC 24150

Qy       1176 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24151 AAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCACCCCAGCA 24210

Qy       1236 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24211 GTATTTTTATTTAAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAA 24270

Qy       1296 GTTACAGATGTCA 1308
              |||||||||||||
Db      24271 GTTACAGATGTCA 24283


RESULT 5
US-10-027-632-100265/c
; Sequence 100265, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100265
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-100265

  Query Match           31.7%;  Score 420;  DB 13;  Length 610;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-202;
  Matches  420;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 396
```

```
Qy          61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         395 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 336

Qy         121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         335 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 276

Qy         181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         275 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 216

Qy         241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         215 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 156

Qy         301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         155 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 96

Qy         361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          95 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 36
```

RESULT 6
US-10-027-632-100265/c
; Sequence 100265, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100265
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-100265

Query Match 31.7%; Score 420; DB 16; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.7e-202;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         455 TCCACAGTGGTCACAGGTAGTACCTGGTCCTAGGGTTGCCTGAGAGCCAACCTCTCCTGC 396

Qy          61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         395 CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT 336

Qy         121 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         335 GATGCTGTGGCCTGTGGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCA 276

Qy         181 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         275 CCACCTTCCTGCCCACAGGCCGAGGGACAGGGTGAGGGTATACCCAAAGCTGATGCAGAG 216

Qy         241 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         215 CCCATTAGCCTAAAAGCAACTGCAGGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 156

Qy         301 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         155 CTCTGAACAAGAGTCCAGCCAACCCTCTTCAGCCAGGCCTCTGTGACCTGCTAGGGTGCA 96

Qy         361 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          95 GGAGGCTTCCAGAAGCAGTTGTTGTAATTAGGACCCAAGCACTGGGAGGGGCTGTTGGCT 36
```

RESULT 7
US-09-908-975-16197
; Sequence 16197, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16197
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-16197

Query Match 4.5%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        1169 CTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCAC 1228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCTTCAACTACCAGAATCTGGGCCAC 60
```

RESULT 8
US-10-116-275-299/c
; Sequence 299, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 299
;   LENGTH: 1652
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-116-275-299

  Query Match             2.2%;  Score 29;  DB 16;  Length 1652;
  Best Local Similarity   100.0%;  Pred. No. 0.00059;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||||||
Db       1474 ACAGATGTCAAAAAAAAAAAAAAAAAAAA 1446


RESULT 9
US-10-437-963-80513
; Sequence 80513, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 80513
;   LENGTH: 1754
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_80128C.1
US-10-437-963-80513

  Query Match             2.0%;  Score 27;  DB 17;  Length 1754;
  Best Local Similarity   100.0%;  Pred. No. 0.0061;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||||
Db       1728 AGATGTCAAAAAAAAAAAAAAAAAAAAA 1754


RESULT 10
US-10-087-192-2014/c
; Sequence 2014, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2014
;   LENGTH: 248436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(248436)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-2014

  Query Match             2.0%;  Score 27;  DB 13;  Length 248436;
  Best Local Similarity   100.0%;  Pred. No. 0.0053;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1301 AGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||||
Db     187916 AGATGTCAAAAAAAAAAAAAAAAAAAAA 187890


RESULT 11
US-09-969-034-4195
; Sequence 4195, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
;  APPLICANT: Burgess, Christopher C.
;  APPLICANT:  Astle, Jon H.
;  APPLICANT:  Carroll, Eddie III
;  APPLICANT:  Catino, Theodore J.
;  APPLICANT:  Dwivedi, Poornima
;  APPLICANT:  Molino, Gary A.
;  APPLICANT:  Thiagalingam, Arunthathi
;  APPLICANT:  Lewis, Marcia E.
;  TITLE OF INVENTION: Nucleic Acid Sequences Differentially
;  TITLE OF INVENTION:  Expressed in Cancer Tissue
;  FILE REFERENCE: 1657/1032
;  CURRENT APPLICATION NUMBER: US/09/969,034
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: 60/237,271
;  PRIOR FILING DATE: 2000-02-10
;  NUMBER OF SEQ ID NOS: 4494
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4195
;   LENGTH: 273
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-034-4195

  Query Match             2.0%;  Score 26;  DB 12;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||
Db        245 GATGTCAAAAAAAAAAAAAAAAAAAA 270


RESULT 12
US-09-960-352-6255/c
; Sequence 6255, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6255
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 27-LIB3057-011-Q1-K1-G3
US-09-960-352-6255
```

```
  Query Match             2.0%;  Score 26;  DB 9;  Length 393;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||
Db         49 GATGTCAAAAAAAAAAAAAAAAAAAA 24


RESULT 13
US-09-960-352-13315/c
; Sequence 13315, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13315
;   LENGTH: 411
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 57-LIB3057-002-Q1-K1-G2
US-09-960-352-13315

  Query Match             2.0%;  Score 26;  DB 9;  Length 411;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||
Db         48 GATGTCAAAAAAAAAAAAAAAAAAAA 23


RESULT 14
US-10-295-027-141
; Sequence 141, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 141
;   LENGTH: 4702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-295-027-141

  Query Match             2.0%;  Score 26;  DB 16;  Length 4702;
  Best Local Similarity   100.0%;  Pred. No. 0.019;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||||
Db       4660 GATGTCAAAAAAAAAAAAAAAAAAAA 4685


RESULT 15
US-10-295-027-1164
; Sequence 1164, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1164
;   LENGTH: 4702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-10-295-027-1164

Query Match 2.0%; Score 26; DB 16; Length 4702;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 4660 GATGTCAAAAAAAAAAAAAAAAAAAA 4685

RESULT 16
US-10-437-963-87284
; Sequence 87284, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87284
; LENGTH: 6808
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(6808)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86244C.1
US-10-437-963-87284

Query Match 2.0%; Score 26; DB 17; Length 6808;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 6768 GATGTCAAAAAAAAAAAAAAAAAAAA 6793

RESULT 17
US-10-085-783A-47925
; Sequence 47925, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47925
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-47925

Query Match 1.9%; Score 25; DB 13; Length 136;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
|||||||||||||||||||||||||
Db 104 ATGTCAAAAAAAAAAAAAAAAAAAA 128

RESULT 18
US-10-242-535A-47925
; Sequence 47925, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47925
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-47925

Query Match 1.9%; Score 25; DB 16; Length 136;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
|||||||||||||||||||||||||
Db 104 ATGTCAAAAAAAAAAAAAAAAAAAA 128

RESULT 19
US-10-085-783A-51657
; Sequence 51657, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51657
; LENGTH: 143
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-51657

  Query Match             1.9%;  Score 25;  DB 13;  Length 143;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        113 ATGTCAAAAAAAAAAAAAAAAAAAA 137


RESULT 20
US-10-242-535A-51657
; Sequence 51657, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51657
;   LENGTH: 143
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-51657

  Query Match             1.9%;  Score 25;  DB 16;  Length 143;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        113 ATGTCAAAAAAAAAAAAAAAAAAAA 137


RESULT 21
US-10-085-783A-45661
; Sequence 45661, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45661
;   LENGTH: 156
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-45661

  Query Match             1.9%;  Score 25;  DB 13;  Length 156;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        125 ATGTCAAAAAAAAAAAAAAAAAAAA 149


RESULT 22
US-10-242-535A-45661
; Sequence 45661, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45661
;   LENGTH: 156
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-45661

  Query Match             1.9%;  Score 25;  DB 16;  Length 156;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        125 ATGTCAAAAAAAAAAAAAAAAAAAA 149


RESULT 23
US-09-867-701-9253
; Sequence 9253, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9253
;   LENGTH: 184
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-9253

  Query Match             1.9%;  Score 25;  DB 9;  Length 184;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
```

```
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        141 ATGTCAAAAAAAAAAAAAAAAAAAA 165


RESULT 24
US-09-962-436-134/c
; Sequence 134, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
;  APPLICANT: Soppet, Daniel
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-75
;  CURRENT APPLICATION NUMBER: US/09/962,436
;  CURRENT FILING DATE:  2001-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,082
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/234,924
;  PRIOR FILING DATE: 2000-09-25
;  NUMBER OF SEQ ID NOS: 568
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 134
;   LENGTH: 191
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-962-436-134

  Query Match             1.9%;  Score 25;  DB 9;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         30 ATGTCAAAAAAAAAAAAAAAAAAAA 6


RESULT 25
US-09-954-456-1020/c
; Sequence 1020, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1020
;   LENGTH: 191
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-1020

  Query Match             1.9%;  Score 25;  DB 9;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         30 ATGTCAAAAAAAAAAAAAAAAAAAA 6


RESULT 26
US-09-880-107-525/c
; Sequence 525, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 525
;   LENGTH: 191
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA243133
US-09-880-107-525

  Query Match             1.9%;  Score 25;  DB 9;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         30 ATGTCAAAAAAAAAAAAAAAAAAAA 6


RESULT 27
US-09-867-701-10779
; Sequence 10779, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10779
;   LENGTH: 207
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-10779

  Query Match             1.9%;  Score 25;  DB 9;  Length 207;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
```

```
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        164 ATGTCAAAAAAAAAAAAAAAAAAAA 188


RESULT 28
US-09-880-107-587/c
; Sequence 587, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 587
;   LENGTH: 222
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA258131
US-09-880-107-587

  Query Match             1.9%;  Score 25;  DB 9;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         39 ATGTCAAAAAAAAAAAAAAAAAAAA 15


RESULT 29
US-10-424-599-31507/c
; Sequence 31507, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 31507
;   LENGTH: 256
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_128456C.1
US-10-424-599-31507

  Query Match             1.9%;  Score 25;  DB 13;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         32 ATGTCAAAAAAAAAAAAAAAAAAAA 8


RESULT 30
US-10-085-783A-52889
; Sequence 52889, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52889
;   LENGTH: 296
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-52889

  Query Match             1.9%;  Score 25;  DB 13;  Length 296;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        265 ATGTCAAAAAAAAAAAAAAAAAAAA 289


RESULT 31
US-10-242-535A-52889
; Sequence 52889, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52889
;   LENGTH: 296
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-52889

  Query Match             1.9%;  Score 25;  DB 16;  Length 296;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        265 ATGTCAAAAAAAAAAAAAAAAAAAA 289


RESULT 32
```

```
US-09-813-358-64/c
; Sequence 64, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Pyle, Ruth
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
;  FILE REFERENCE: 210121.501
;  CURRENT APPLICATION NUMBER: US/09/813,358
;  CURRENT FILING DATE:  2001-03-21
;  NUMBER OF SEQ ID NOS: 222
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
;   LENGTH: 330
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(330)
;   OTHER INFORMATION: n = A,T,C or G
US-09-813-358-64

  Query Match             1.9%;  Score 25;  DB 9;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         30 ATGTCAAAAAAAAAAAAAAAAAAAA 6


RESULT 33
US-09-997-279-64/c
; Sequence 64, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
;  APPLICANT: Chenault, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
;  FILE REFERENCE: 210121.501C1
;  CURRENT APPLICATION NUMBER: US/09/997,279
;  CURRENT FILING DATE:  2001-11-28
;  NUMBER OF SEQ ID NOS: 230
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
;   LENGTH: 330
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(330)
;   OTHER INFORMATION: n = A,T,C or G
US-09-997-279-64

  Query Match             1.9%;  Score 25;  DB 10;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         30 ATGTCAAAAAAAAAAAAAAAAAAAA 6


RESULT 34
US-09-922-293-42/c
; Sequence 42, Application US/09922293
; Publication No. US20040123339A1
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Heck, Gregory R.
;  APPLICANT:  Liu, Jingdong
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Transcription in Plants
;  FILE REFERENCE: 16517.254
;  CURRENT APPLICATION NUMBER: US/09/922,293
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: US 60/067,000
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: US 60/069,472
;  PRIOR FILING DATE: 1997-12-09
;  PRIOR APPLICATION NUMBER: US 60/071,479
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/074,201
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,282
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,280
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,281
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,566
;  PRIOR FILING DATE: 1998-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,567
;  PRIOR FILING DATE: 1998-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,565
;  PRIOR FILING DATE: 1998-02-12
;  PRIOR APPLICATION NUMBER: US 60/075,462
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/074,789
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,459
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,461
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,464
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,460
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,463
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/077,231
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 60/077,229
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 60/077,230
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 60/078,368
;  PRIOR FILING DATE: 1998-03-18
;  PRIOR APPLICATION NUMBER: US 60/080,844
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: US 60/083,067
;  PRIOR FILING DATE: 1998-04-27
;  PRIOR APPLICATION NUMBER: US 60/083,386
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/083,387
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/083,388
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/083,389
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/085,224
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,223
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,222
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,533
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: US 60/086,186
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,187
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,185
```

; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,184
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,183
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,188
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/089,524
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: US 60/089,810
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,814
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,808
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,812
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,807
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,806
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,813
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,811
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,793
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/091,405
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/091,247
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/099,667
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/099,668
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/099,670
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/099,697
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/100,674
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 60/100,673
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 60/100,672
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 60/100,963
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 60/101,131
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 60/101,132
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 60/101,130
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 60/101,508
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,344
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,347
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,343
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,707
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/104,126
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,128
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,127
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,124
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/109,018
; PRIOR FILING DATE: 1998-11-18

; PRIOR APPLICATION NUMBER: US 60/108,996
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: US 09/199,129
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: US 09/210,297
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 60/111,981
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/113,224
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/229,413
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 3853
; SEQ ID NO 42
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-922-293-42

Query Match 1.9%; Score 25; DB 12; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 101 ATGTCAAAAAAAAAAAAAAAAAAAA 77

RESULT 35
US-09-960-352-1456/c
; Sequence 1456, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1456
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 07-LIB3057-005-Q1-K1-B3
US-09-960-352-1456

Query Match 1.9%; Score 25; DB 9; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1302 GATGTCAAAAAAAAAAAAAAAAAAA 1326
|||||||||||||||||||||||||
Db 39 GATGTCAAAAAAAAAAAAAAAAAAA 15

RESULT 36
US-10-424-599-52765
; Sequence 52765, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

```
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 52765
;   LENGTH: 358
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_1865C.1
US-10-424-599-52765

  Query Match             1.9%;  Score 25;  DB 13;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 0.067;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         22 ATGTCAAAAAAAAAAAAAAAAAAAA 46


RESULT 37
US-10-437-963-88420
; Sequence 88420, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 88420
;   LENGTH: 382
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(382)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_87273C.1
US-10-437-963-88420

  Query Match             1.9%;  Score 25;  DB 17;  Length 382;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         355 ATGTCAAAAAAAAAAAAAAAAAAAA 379


RESULT 38
US-10-175-184A-31
; Sequence 31, Application US/10175184A
; Publication No. US20040038292A1
; GENERAL INFORMATION:
;  APPLICANT: Pfizer Inc.
;  APPLICANT:  Burslem, Martyn Frank
;  APPLICANT:  Johnson, Claire Michelle
;  APPLICANT:  Cooper, Lisa
;  APPLICANT:  Martin, Paul
;  TITLE OF INVENTION: WOUND HEALING BIOMARKERS
;  FILE REFERENCE: PC22024AGLK
;  CURRENT APPLICATION NUMBER: US/10/175,184A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 60/305,346
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: GB 0114869.1
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
;   LENGTH: 390
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-175-184A-31

  Query Match             1.9%;  Score 25;  DB 13;  Length 390;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         356 ATGTCAAAAAAAAAAAAAAAAAAAA 380


RESULT 39
US-09-918-995-7699
; Sequence 7699, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7699
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-7699

  Query Match             1.9%;  Score 25;  DB 10;  Length 430;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         405 ATGTCAAAAAAAAAAAAAAAAAAAA 429


RESULT 40
US-09-960-352-8587/c
; Sequence 8587, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8587
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 37-LIB3057-005-Q1-K1-B2
US-09-960-352-8587

  Query Match             1.9%;  Score 25;  DB 9;  Length 464;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GATGTCAAAAAAAAAAAAAAAAAAA 1326
              |||||||||||||||||||||||||
Db         42 GATGTCAAAAAAAAAAAAAAAAAAA 18


RESULT 41
US-10-198-846-1221/c
; Sequence 1221, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1221
;   LENGTH: 467
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 5, 7, 14, 116, 200, 230, 243, 245, 252, 294, 312, 340, 349,
;   LOCATION: 355, 446, 448, 452, 462
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1221

  Query Match             1.9%;  Score 25;  DB 15;  Length 467;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         99 ATGTCAAAAAAAAAAAAAAAAAAA 75


RESULT 42
US-10-102-524-36/c
; Sequence 36, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Gordon, Brian
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF KIDNEY CANCER
;  FILE REFERENCE: 210121.572
;  CURRENT APPLICATION NUMBER: US/10/102,524
;  CURRENT FILING DATE:  2002-03-19
;  NUMBER OF SEQ ID NOS: 1863
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 64, 100, 111, 120, 148, 163, 298, 327, 338, 430, 455, 462,
;   LOCATION: 471
;   OTHER INFORMATION: n = A,T,C or G
US-10-102-524-36

  Query Match             1.9%;  Score 25;  DB 15;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db         39 ATGTCAAAAAAAAAAAAAAAAAAA 15


RESULT 43
US-10-437-963-85418
; Sequence 85418, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 85418
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(483)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_84559C.1
US-10-437-963-85418

  Query Match             1.9%;  Score 25;  DB 17;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        421 ATGTCAAAAAAAAAAAAAAAAAAA 445


RESULT 44
US-10-027-632-95705
; Sequence 95705, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
```

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95705
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-95705

Query Match 1.9%; Score 25; DB 13; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 143 ATGTCAAAAAAAAAAAAAAAAAAAA 167

RESULT 45
US-10-027-632-305902
; Sequence 305902, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 305902
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305902

Query Match 1.9%; Score 25; DB 13; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 143 ATGTCAAAAAAAAAAAAAAAAAAAA 167

RESULT 46
US-10-027-632-95705
; Sequence 95705, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95705
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-95705

Query Match 1.9%; Score 25; DB 16; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 143 ATGTCAAAAAAAAAAAAAAAAAAAA 167

RESULT 47
US-10-027-632-305902
; Sequence 305902, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 305902
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305902

Query Match 1.9%; Score 25; DB 16; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||||
Db 143 ATGTCAAAAAAAAAAAAAAAAAAAA 167

```
RESULT 48
US-10-437-963-19508/c
; Sequence 19508, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 19508
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_24962C.1
US-10-437-963-19508

  Query Match             1.9%;  Score 25;  DB 17;  Length 492;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db         35 ATGTCAAAAAAAAAAAAAAAAAAAA 11


RESULT 49
US-10-029-386-1201/c
; Sequence 1201, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1201
;   LENGTH: 503
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR22_30.0
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: NT HIT: M23442.1, EVALUE 3.10e-02
US-10-029-386-1201

  Query Match             1.9%;  Score 25;  DB 15;  Length 503;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db        230 ATGTCAAAAAAAAAAAAAAAAAAAA 206


RESULT 50
US-09-925-301-787
; Sequence 787, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 787
;   LENGTH: 527
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (492)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-787

  Query Match             1.9%;  Score 25;  DB 9;  Length 527;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db        501 ATGTCAAAAAAAAAAAAAAAAAAAA 525


Search completed: September 17, 2004, 06:45:02
Job time : 3730 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 02:40:43 ; Search time 123 Seconds
(without alignments)
5987.153 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtggtcacaggtag.....aaaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	25	1.9	469	1	US-08-468-347-23	Sequence 23, Appl
	2	25	1.9	469	1	US-08-226-264-25	Sequence 25, Appl
	3	25	1.9	469	1	US-08-467-389-23	Sequence 23, Appl
	4	25	1.9	469	2	US-08-779-379-23	Sequence 23, Appl
	5	25	1.9	469	2	US-08-469-219-23	Sequence 23, Appl
	6	25	1.9	469	3	US-09-228-152-23	Sequence 23, Appl
	7	25	1.9	729	1	US-08-447-010-1	Sequence 1, Appli
	8	25	1.9	868	3	US-08-889-502-20	Sequence 20, Appl
c	9	25	1.9	98844	4	US-09-791-211-10	Sequence 10, Appl
c	10	24	1.8	40	3	US-09-306-290-26	Sequence 26, Appl
c	11	24	1.8	555	4	US-09-449-285A-15	Sequence 15, Appl
	12	24	1.8	1196	3	US-07-959-509-4	Sequence 4, Appli
	13	24	1.8	1332	3	US-09-333-423-1	Sequence 1, Appli
	14	24	1.8	1392	4	US-08-957-351-1	Sequence 1, Appli
	15	24	1.8	1478	4	US-09-216-393B-7	Sequence 7, Appli
	16	24	1.8	1496	4	US-09-712-529-1	Sequence 1, Appli
	17	24	1.8	1628	2	US-08-883-515-3	Sequence 3, Appli
	18	24	1.8	1705	4	US-09-205-258-216	Sequence 216, App
	19	24	1.8	2017	3	US-09-436-983-1	Sequence 1, Appli
	20	24	1.8	2193	4	US-09-427-261-2	Sequence 2, Appli
	21	24	1.8	2193	4	US-09-427-261-3	Sequence 3, Appli
	22	24	1.8	3001	4	US-09-539-333D-222	Sequence 222, App
	23	24	1.8	3848	3	US-09-112-096-28	Sequence 28, Appl
c	24	24	1.8	3952	2	US-08-381-691-16	Sequence 16, Appl
	25	24	1.8	5668	3	US-09-112-096-14	Sequence 14, Appl
	26	24	1.8	5668	4	US-09-636-215-777	Sequence 777, App
	27	24	1.8	5668	4	US-09-685-166A-777	Sequence 777, App
c	28	24	1.8	21234	4	US-09-810-671-3	Sequence 3, Appli
c	29	24	1.8	21234	4	US-10-109-854-3	Sequence 3, Appli
c	30	24	1.8	32042	4	US-09-245-281-44	Sequence 44, Appl
c	31	24	1.8	32042	4	US-09-340-620A-63	Sequence 63, Appl
c	32	24	1.8	50000	4	US-09-146-053-3	Sequence 3, Appli
c	33	24	1.8	786431	4	US-09-751-389-3	Sequence 3, Appli
c	34	23	1.7	40	3	US-09-306-290-15	Sequence 15, Appl
	35	23	1.7	67	4	US-09-621-976-10956	Sequence 10956, A
	36	23	1.7	97	4	US-09-621-976-9086	Sequence 9086, Ap
	37	23	1.7	307	4	US-09-091-725-28	Sequence 28, Appl
	38	23	1.7	350	4	US-09-621-976-15342	Sequence 15342, A
	39	23	1.7	433	1	US-07-987-272A-13	Sequence 13, Appl
	40	23	1.7	567	1	US-08-661-168-6	Sequence 6, Appli
	41	23	1.7	627	3	US-09-385-982-4	Sequence 4, Appli
	42	23	1.7	989	2	US-08-874-460-1	Sequence 1, Appli
	43	23	1.7	989	4	US-09-272-162-1	Sequence 1, Appli
	44	23	1.7	1005	4	US-09-647-224A-5	Sequence 5, Appli
	45	23	1.7	1087	4	US-09-396-149-17	Sequence 17, Appl
	46	23	1.7	1108	1	US-08-036-555B-135	Sequence 135, App
	47	23	1.7	1108	1	US-08-469-569-135	Sequence 135, App
	48	23	1.7	1108	1	US-08-249-322A-135	Sequence 135, App
	49	23	1.7	1108	1	US-08-469-526A-135	Sequence 135, App
	50	23	1.7	1108	2	US-08-734-591A-135	Sequence 135, App
	51	23	1.7	1108	2	US-08-469-660-135	Sequence 135, App
	52	23	1.7	1108	3	US-08-341-018-5	Sequence 5, Appli
	53	23	1.7	1108	3	US-08-470-335-135	Sequence 135, App
	54	23	1.7	1108	3	US-08-735-021-135	Sequence 135, App
	55	23	1.7	1108	3	US-08-734-664A-135	Sequence 135, App
	56	23	1.7	1108	3	US-08-470-339-135	Sequence 135, App
	57	23	1.7	1108	4	US-08-467-602-135	Sequence 135, App
	58	23	1.7	1108	5	PCT-US94-05083C-131	Sequence 131, App
	59	23	1.7	1108	5	PCT-US95-06846A-135	Sequence 135, App
	60	23	1.7	1193	1	US-08-036-555B-134	Sequence 134, App
	61	23	1.7	1193	1	US-08-469-569-134	Sequence 134, App
	62	23	1.7	1193	1	US-08-249-322A-134	Sequence 134, App
	63	23	1.7	1193	1	US-08-469-526A-134	Sequence 134, App
	64	23	1.7	1193	2	US-08-734-591A-134	Sequence 134, App
	65	23	1.7	1193	2	US-08-469-660-134	Sequence 134, App
	66	23	1.7	1193	3	US-08-341-018-3	Sequence 3, Appli
	67	23	1.7	1193	3	US-08-470-335-134	Sequence 134, App
	68	23	1.7	1193	3	US-08-735-021-134	Sequence 134, App
	69	23	1.7	1193	3	US-08-734-664A-134	Sequence 134, App
	70	23	1.7	1193	3	US-08-470-339-134	Sequence 134, App
	71	23	1.7	1193	4	US-08-467-602-134	Sequence 134, App
	72	23	1.7	1193	5	PCT-US94-05083C-130	Sequence 130, App
	73	23	1.7	1193	5	PCT-US95-06846A-134	Sequence 134, App
	74	23	1.7	1209	4	US-09-614-912-61	Sequence 61, Appl
	75	23	1.7	1291	4	US-08-555-755C-5	Sequence 5, Appli
	76	23	1.7	1364	2	US-08-815-718-1	Sequence 1, Appli
	77	23	1.7	1509	4	US-09-149-476-173	Sequence 173, App
	78	23	1.7	1559	4	US-09-482-273-56	Sequence 56, Appl
	79	23	1.7	1575	4	US-09-833-381-1159	Sequence 1159, Ap
c	80	23	1.7	1703	3	US-09-135-021-77	Sequence 77, Appl
c	81	23	1.7	1703	3	US-09-135-020-3	Sequence 3, Appli
c	82	23	1.7	1703	3	US-09-135-010A-3	Sequence 3, Appli
c	83	23	1.7	1703	4	US-09-444-871-3	Sequence 3, Appli
c	84	23	1.7	1703	4	US-09-597-735-3	Sequence 3, Appli
c	85	23	1.7	1703	4	US-09-444-295-3	Sequence 3, Appli
c	86	23	1.7	1703	4	US-09-597-732-3	Sequence 3, Appli
c	87	23	1.7	1703	4	US-09-597-731-3	Sequence 3, Appli
	88	23	1.7	1741	1	US-08-565-655-5	Sequence 5, Appli
	89	23	1.7	1769	4	US-09-333-593A-5	Sequence 5, Appli
	90	23	1.7	1769	4	US-09-505-250-1	Sequence 1, Appli
	91	23	1.7	1769	4	US-09-016-434-1072	Sequence 1072, Ap
	92	23	1.7	1769	4	US-09-023-655-888	Sequence 888, App
c	93	23	1.7	1821	3	US-09-056-105-23	Sequence 23, Appl
c	94	23	1.7	1896	1	US-08-253-503-1	Sequence 1, Appli
c	95	23	1.7	1896	1	US-08-796-883-1	Sequence 1, Appli
c	96	23	1.7	1896	2	US-08-611-273B-1	Sequence 1, Appli
c	97	23	1.7	1896	2	US-08-531-864-1	Sequence 1, Appli
c	98	23	1.7	1896	2	US-08-373-636C-1	Sequence 1, Appli
c	99	23	1.7	1896	3	US-08-602-506A-1	Sequence 1, Appli
c	100	23	1.7	1896	3	US-09-266-294-1	Sequence 1, Appli

c 101	23	1.7	1896	3	US-09-179-281-1	Sequence 1, Appli
c 102	23	1.7	1896	3	US-09-056-105-25	Sequence 25, Appl
103	23	1.7	1969	2	US-08-541-033A-7	Sequence 7, Appli
104	23	1.7	1969	2	US-08-828-451-7	Sequence 7, Appli
105	23	1.7	2026	1	US-08-503-133A-1	Sequence 1, Appli
106	23	1.7	2026	2	US-08-576-775A-1	Sequence 1, Appli
107	23	1.7	2026	2	US-08-972-498-1	Sequence 1, Appli
108	23	1.7	2026	3	US-08-899-545-1	Sequence 1, Appli
109	23	1.7	2053	4	US-09-833-381-1403	Sequence 1403, Ap
110	23	1.7	2096	2	US-08-541-033A-19	Sequence 19, Appl
111	23	1.7	2096	2	US-08-828-451-19	Sequence 19, Appl
112	23	1.7	2099	2	US-08-541-033A-3	Sequence 3, Appli
113	23	1.7	2099	2	US-08-828-451-3	Sequence 3, Appli
114	23	1.7	2103	2	US-08-897-340-2	Sequence 2, Appli
115	23	1.7	2103	3	US-09-252-329-2	Sequence 2, Appli
116	23	1.7	2119	2	US-08-381-691-17	Sequence 17, Appl
117	23	1.7	2137	2	US-08-541-033A-18	Sequence 18, Appl
118	23	1.7	2137	2	US-08-828-451-18	Sequence 18, Appl
119	23	1.7	2140	2	US-08-541-033A-1	Sequence 1, Appli
120	23	1.7	2140	2	US-08-828-451-1	Sequence 1, Appli
121	23	1.7	2198	2	US-08-755-728-2	Sequence 2, Appli
122	23	1.7	2198	2	US-08-974-655-2	Sequence 2, Appli
123	23	1.7	2198	3	US-09-283-011-2	Sequence 2, Appli
124	23	1.7	2209	4	US-09-807-258-25	Sequence 25, Appl
125	23	1.7	2335	3	US-09-387-574-9	Sequence 9, Appli
126	23	1.7	2335	4	US-09-668-096-9	Sequence 9, Appli
127	23	1.7	2403	4	US-09-614-912-99	Sequence 99, Appl
128	23	1.7	2420	5	PCT-US93-00227-1	Sequence 1, Appli
129	23	1.7	2759	4	US-09-144-367-1	Sequence 1, Appli
c 130	23	1.7	2929	4	US-09-705-299-10	Sequence 10, Appl
131	23	1.7	3238	3	US-08-123-934A-5	Sequence 5, Appli
132	23	1.7	3238	4	US-09-874-628-5	Sequence 5, Appli
133	23	1.7	3238	5	PCT-US94-10080-5	Sequence 5, Appli
134	23	1.7	3268	4	US-09-688-188B-1	Sequence 1, Appli
135	23	1.7	3268	4	US-09-291-417D-1	Sequence 1, Appli
136	23	1.7	3300	4	US-09-336-643A-82	Sequence 82, Appl
137	23	1.7	3377	4	US-09-819-989-1	Sequence 1, Appli
138	23	1.7	3377	4	US-10-273-992-1	Sequence 1, Appli
139	23	1.7	3441	4	US-09-866-028-6	Sequence 6, Appli
140	23	1.7	3946	3	US-09-083-351-1	Sequence 1, Appli
141	23	1.7	3946	3	US-09-083-352-1	Sequence 1, Appli
142	23	1.7	4113	4	US-09-785-381-2	Sequence 2, Appli
143	23	1.7	4843	3	US-08-986-485-1	Sequence 1, Appli
144	23	1.7	4874	4	US-09-187-330-2	Sequence 2, Appli
c 145	23	1.7	6182	4	US-10-204-708-88	Sequence 88, Appl
c 146	23	1.7	6405	4	US-09-281-481A-18	Sequence 18, Appl
147	23	1.7	6644	4	US-08-875-435B-5	Sequence 5, Appli
148	23	1.7	8201	1	US-08-253-155A-9	Sequence 9, Appli
149	23	1.7	9278	1	US-08-243-542-9	Sequence 9, Appli
150	23	1.7	9278	1	US-08-477-407-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-08-468-347-23
; Sequence 23, Application US/08468347
; Patent No. 5783421
;  GENERAL INFORMATION:
;    APPLICANT:  Zeelon, Elisha P.
;    APPLICANT:  Werber, Moshe M.
;    APPLICANT:  Levanon, Avigdor
;    TITLE OF INVENTION:  NOVEL POLYPEPTIDE HAVING FACTOR Xa
;    TITLE OF INVENTION:  INHIBITORY ACTIVITY
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,347
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/225,442
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0317/43020-A/JPW/EAB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-977-9550
;      TELEFAX:  212-664-0525
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  469 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-468-347-23

  Query Match             1.9%;  Score 25;  DB 1;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 2
US-08-226-264-25
; Sequence 25, Application US/08226264
; Patent No. 5801017
;  GENERAL INFORMATION:
;    APPLICANT:  Werber, Moshe M.
;    APPLICANT:  Zeelon, Elisha P.
;    APPLICANT:  Levanon, Avigdor
;    APPLICANT:  Guy, Rachel
;    APPLICANT:  Goldlust, Arie
;    APPLICANT:  Rigbi, Meir
;    APPLICANT:  Panet, Amos
;    APPLICANT:  Fischer, Meir
;    TITLE OF INVENTION:  PRODUCTION OF RECOMBINANT FACTOR XA
;    TITLE OF INVENTION:  INHIBITORS OF LEECH HIRUDO MEDICINALIS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/226,264
;      FILING DATE:  08-APR-94
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
```

```
;       REGISTRATION NUMBER:  28,678
;       REFERENCE/DOCKET NUMBER:  40017-A/JPW/GJG
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-278-0400
;       TELEFAX:  212-391-0525
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  469 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-226-264-25

  Query Match             1.9%;  Score 25;  DB 1;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 3
US-08-467-389-23
; Sequence 23, Application US/08467389
; Patent No. 5824641
;  GENERAL INFORMATION:
;     APPLICANT:  Zeelon, Elisha P.
;     APPLICANT:  Werber, Moshe M.
;     APPLICANT:  Levanon, Avigdor
;     TITLE OF INVENTION:  NOVEL POLYPEPTIDE HAVING FACTOR Xa
;     TITLE OF INVENTION:  INHIBITORY ACTIVITY
;     NUMBER OF SEQUENCES:  25
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cooper & Dunham
;       STREET:  30 Rockefeller Plaza
;       CITY:  New York
;       STATE:  New York
;       ZIP:  10112
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/467,389
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/225,442
;       FILING DATE:  08-APR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  White, John P.
;       REGISTRATION NUMBER:  28,678
;       REFERENCE/DOCKET NUMBER:  0317/43020-A/JPW/EAB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-977-9550
;       TELEFAX:  212-664-0525
;   INFORMATION FOR SEQ ID NO:  23:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  469 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
US-08-467-389-23

  Query Match             1.9%;  Score 25;  DB 1;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 4
US-08-779-379-23
; Sequence 23, Application US/08779379
; Patent No. 5858970
;  GENERAL INFORMATION:
;     APPLICANT:  Zeelon, Elisha P.
;     APPLICANT:  Werber, Moshe M.
;     APPLICANT:  Levanon, Avigdor
;     TITLE OF INVENTION:  NOVEL POLYPEPTIDE HAVING FACTOR Xa
;     TITLE OF INVENTION:  INHIBITORY ACTIVITY
;     NUMBER OF SEQUENCES:  25
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cooper & Dunham
;       STREET:  30 Rockefeller Plaza
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10112
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/779,379
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/225,442
;       FILING DATE:  08-APR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  White, John P.
;       REGISTRATION NUMBER:  28,678
;       REFERENCE/DOCKET NUMBER:  0317/43020-A/JPW/EAB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-977-9550
;       TELEFAX:  212-664-0525
;   INFORMATION FOR SEQ ID NO:  23:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  469 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
US-08-779-379-23

  Query Match             1.9%;  Score 25;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 5
US-08-469-219-23
; Sequence 23, Application US/08469219
; Patent No. 5863534
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Zeelon, Elisha P.
;    APPLICANT:  Werber, Moshe M.
;    APPLICANT:  Levanon, Avigdor
;    TITLE OF INVENTION:  NOVEL POLYPEPTIDE HAVING FACTOR Xa
;    TITLE OF INVENTION:  INHIBITORY ACTIVITY
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,219
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/225,442
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0317/43020-A/JPW/EAB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-977-9550
;      TELEFAX:  212-664-0525
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  469 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-469-219-23

  Query Match             1.9%;  Score 25;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 6
US-09-228-152-23
; Sequence 23, Application US/09228152
; Patent No. 6211341
; GENERAL INFORMATION:
;  APPLICANT: Zeelon, Elisha P.
;  APPLICANT: Werber, Moshe M.
;  APPLICANT: Levanon, Avigdor
;  TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa INHIBITORY ACTIVITY
;  FILE REFERENCE: 43020aya
;  CURRENT APPLICATION NUMBER: US/09/228,152
;  CURRENT FILING DATE: 1999-01-11
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;    LENGTH: 469
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:  DNA sequence
;    OTHER INFORMATION: of clone pSP65-XaI-4.
US-09-228-152-23

  Query Match             1.9%;  Score 25;  DB 3;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db        445 ATGTCAAAAAAAAAAAAAAAAAAAA 469


RESULT 7
US-08-447-010-1
; Sequence 1, Application US/08447010
; Patent No. 5770718
;  GENERAL INFORMATION:
;    APPLICANT:  MOFFATT, BARBARA
;    TITLE OF INVENTION:  GENE FOR APRT FROM PLANT TISSUE
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  330 University Avenue, Suite 701
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,010
;      FILING DATE:  22-MAY-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/230,695
;      FILING DATE:  21-APR-1994
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/888,132
;      FILING DATE:  26-MAY-1992
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEWART, MICHAEL I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1811-183 MIS:vg
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;      TELEX:  065-24567 SIMBAS
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  729 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  both
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(18..569)
US-08-447-010-1

  Query Match             1.9%;  Score 25;  DB 1;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
             |||||||||||||||||||||||||
Db        705 ATGTCAAAAAAAAAAAAAAAAAAAA 729


RESULT 8
```

```
US-08-889-502-20
; Sequence 20, Application US/08889502
; Patent No. 6066726
;  GENERAL INFORMATION:
;    APPLICANT:  Farb, David H
;    APPLICANT:  Russek, Shelley J
;    TITLE OF INVENTION:  GENE THERAPY VECTOR WITH TISSUE
;    TITLE OF INVENTION:  SPECIFICITY
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kevin M. Farrell
;      STREET:  P.O. Box 999
;      CITY:  York Harbor
;      STATE:  ME
;      COUNTRY:  USA
;      ZIP:  03911
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/889,502
;      FILING DATE:  08-JUL-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Farrell, Kevin M
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  0146-2008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (207) 363-0558
;      TELEFAX:  (207) 363-0528
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  868 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-889-502-20

  Query Match             1.9%;  Score 25;  DB 3;  Length 868;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db        440 ATGTCAAAAAAAAAAAAAAAAAAAA 464


RESULT 9
US-09-791-211-10/c
; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
;  APPLICANT:  Donna T. Ward
;  APPLICANT:  Andrew T. Watt
;  TITLE OF INVENTION:  ANTISENSE MODULATION OF WRN EXPRESSION
;  FILE REFERENCE:  RTS-0205
;  CURRENT APPLICATION NUMBER: US/09/791,211
;  CURRENT FILING DATE:  2001-02-23
;  NUMBER OF SEQ ID NOS:  90
; SEQ ID NO 10
;   LENGTH: 98844
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 24962
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 64383
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65468
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65469
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65470
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 65471
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 87130
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 89049
;   OTHER INFORMATION: unknown
;   OTHER INFORMATION:
US-09-791-211-10

  Query Match             1.9%;  Score 25;  DB 4;  Length 98844;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||||
Db      95450 ATGTCAAAAAAAAAAAAAAAAAAAA 95426


RESULT 10
US-09-306-290-26/c
; Sequence 26, Application US/09306290
; Patent No. 6221635
; GENERAL INFORMATION:
;  APPLICANT: Rovera, Giovanni
;  APPLICANT: Mukhopadhyay, Sunil
;  TITLE OF INVENTION: METHODS FOR SOLID-PHASE AMPLIFICATION OF DNA TEMPLATE
;  TITLE OF INVENTION: (SPADT) USING MULTIARRAYS
;  FILE REFERENCE: 09924-10
;  CURRENT APPLICATION NUMBER: US/09/306,290
;  CURRENT FILING DATE: 1999-05-06
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer p41
;   OTHER INFORMATION: FH440
US-09-306-290-26

  Query Match             1.8%;  Score 24;  DB 3;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 0.16;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db         24 TGTCAAAAAAAAAAAAAAAAAAAA 1


RESULT 11
US-09-449-285A-15/c
; Sequence 15, Application US/09449285A
; Patent No. 6313280
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnology
;  TITLE OF INVENTION: SMAD-INTERACTING POLYPEPTIDES AND THEIR USE
;  FILE REFERENCE: 2676-4232US
;  CURRENT APPLICATION NUMBER: US/09/449,285A
;  CURRENT FILING DATE:  1999-11-24
;  PRIOR APPLICATION NUMBER: PCT/EP98/03193
```

```
;  PRIOR FILING DATE: 1998-05-28
;  PRIOR APPLICATION NUMBER: 97201645.5
;  PRIOR FILING DATE: 1997-06-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 555
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (382)..(555)
;   OTHER INFORMATION: n can be any nucleotide
US-09-449-285A-15

  Query Match             1.8%;  Score 24;  DB 4;  Length 555;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db         24 TGTCAAAAAAAAAAAAAAAAAAAA 1

RESULT 12
US-07-959-509-4
; Sequence 4, Application US/07959509
; Patent No. 6001560
;  GENERAL INFORMATION:
;    APPLICANT:  Lonial, Herinder
;    APPLICANT:  Narula, Satwant
;    APPLICANT:  Zavodny, Paul
;    TITLE OF INVENTION:  Human Gamma Interferon Antagonist/Agonist Screen
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  One Giralda Farms
;      CITY:  Madison
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07940
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh 6.0.5
;      SOFTWARE:  Microsoft Word 4.00B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/959,509
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/616,621
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dulak, No. 6001560man C.
;      REGISTRATION NUMBER:  31,608
;      REFERENCE/DOCKET NUMBER:  JB0166
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 822 7375
;      TELEFAX:  201 822 7039
;      TELEX:  219165
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1196 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-07-959-509-4

  Query Match             1.8%;  Score 24;  DB 3;  Length 1196;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db        327 TGTCAAAAAAAAAAAAAAAAAAAA 350

RESULT 13
US-09-333-423-1
; Sequence 1, Application US/09333423
; Patent No. 6265636
; GENERAL INFORMATION:
;  APPLICANT: Randall, Douglas
;  APPLICANT: Thelen, Jay
;  APPLICANT: Miernyk, Jan
;  APPLICANT: Muszynski, Michael
;  APPLICANT: Sewalt, Vincent
;  TITLE OF INVENTION: Pyruvate Dehydrogenase Kinase
;  TITLE OF INVENTION: Polynucleotides, Polypeptides and Uses Thereof
;  FILE REFERENCE: 0818
;  CURRENT APPLICATION NUMBER: US/09/333,423
;  CURRENT FILING DATE: 1999-06-15
;  EARLIER APPLICATION NUMBER: 60/089,998
;  EARLIER FILING DATE: 1998-06-19
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1332
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (55)...(1095)
US-09-333-423-1

  Query Match             1.8%;  Score 24;  DB 3;  Length 1332;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       1279 TGTCAAAAAAAAAAAAAAAAAAAA 1302

RESULT 14
US-08-957-351-1
; Sequence 1, Application US/08957351
; Patent No. 6306586
;  GENERAL INFORMATION:
;    APPLICANT:  Semina, Elena
;    APPLICANT:  Murray, Jeffrey C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR THE
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF CATARACTS
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109-2170
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/957,351
;      FILING DATE:  24-OCT-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Arnold, Beth E.
;      REGISTRATION NUMBER:  35,430
;      REFERENCE/DOCKET NUMBER:  UIA-024.01
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  617-832-1000
;       TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  1392 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-957-351-1

  Query Match             1.8%;  Score 24;  DB 4;  Length 1392;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       1369 TGTCAAAAAAAAAAAAAAAAAAAA 1392


RESULT 15
US-09-216-393B-7
; Sequence 7, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
;  APPLICANT: Milhausen, Michael James
;  TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THER
;  FILE REFERENCE: TX-1-C2
;  CURRENT APPLICATION NUMBER: US/09/216,393B
;  CURRENT FILING DATE:  1998-12-18
;  PRIOR APPLICATION NUMBER: 08/994,825
;  PRIOR FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 366
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 1478
;   TYPE: DNA
;   ORGANISM: Toxoplasma gondii
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (19)..(1161)
;   OTHER INFORMATION:
US-09-216-393B-7

  Query Match             1.8%;  Score 24;  DB 4;  Length 1478;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       1427 TGTCAAAAAAAAAAAAAAAAAAAA 1450


RESULT 16
US-09-712-529-1
; Sequence 1, Application US/09712529
; Patent No. 6485938
; GENERAL INFORMATION:
;  APPLICANT:  Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  APPLICANT:  Whitmore, Theodore E.
;  APPLICANT:  Thompson, Penny P.
;  TITLE OF INVENTION:  Human Zven Proteins
;  FILE REFERENCE:  99-81
;  CURRENT APPLICATION NUMBER: US/09/712,529
;  CURRENT FILING DATE:  2000-11-14
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1496
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (66)...(389)
US-09-712-529-1

  Query Match             1.8%;  Score 24;  DB 4;  Length 1496;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       1459 TGTCAAAAAAAAAAAAAAAAAAAA 1482


RESULT 17
US-08-883-515-3
; Sequence 3, Application US/08883515
; Patent No. 5981836
;  GENERAL INFORMATION:
;    APPLICANT:  Osteryoung, Katherine W
;    TITLE OF INVENTION:  PLANT CHLOROPLAST DIVISION GENES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,515
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J.
;      REGISTRATION NUMBER:  27,386
;      REFERENCE/DOCKET NUMBER:  920905.90016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1628 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1316
US-08-883-515-3

  Query Match             1.8%;  Score 24;  DB 2;  Length 1628;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       1604 TGTCAAAAAAAAAAAAAAAAAAAA 1627


RESULT 18
US-09-205-258-216
; Sequence 216, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
```

; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 216
; LENGTH: 1705
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1281)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1704)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-216

Query Match 1.8%; Score 24; DB 4; Length 1705;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||
Db 1647 TGTCAAAAAAAAAAAAAAAAAAAA 1670

RESULT 19
US-09-436-983-1
; Sequence 1, Application US/09436983
; Patent No. 6294343
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C
; APPLICANT: Wilson, Keith E
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER,
; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF SCREENING FOR COLORECTAL
; TITLE OF INVENTION: CANCER MODULATORS
; FILE REFERENCE: A-68431/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/436,983
; CURRENT FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-436-983-1

Query Match 1.8%; Score 24; DB 3; Length 2017;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
||||||||||||||||||||||||
Db 1981 TGTCAAAAAAAAAAAAAAAAAAAA 2004

RESULT 20
US-09-427-261-2
; Sequence 2, Application US/09427261A
; Patent No. 6368811
; GENERAL INFORMATION:
; APPLICANT: Grootjans, Jan
; APPLICANT: Zimmerman, Pascale
; APPLICANT: David, Guido

```
;  TITLE OF INVENTION: Syndecan Interacting Proteins and the Use Thereof
;  FILE REFERENCE: 2676-4206US
;  CURRENT APPLICATION NUMBER: US/09/427,261A
;  CURRENT FILING DATE: 1999-10-25
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 2193
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:Syndecan DNA
US-09-427-261-2

  Query Match             1.8%;  Score 24;  DB 4;  Length 2193;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       2159 TGTCAAAAAAAAAAAAAAAAAAAA 2182


RESULT 21
US-09-427-261-3
; Sequence 3, Application US/09427261A
; Patent No. 6368811
; GENERAL INFORMATION:
;  APPLICANT: Grootjans, Jan
;  APPLICANT: Zimmerman, Pascale
;  APPLICANT: David, Guido
;  TITLE OF INVENTION: Syndecan Interacting Proteins and the Use Thereof
;  FILE REFERENCE: 2676-4206US
;  CURRENT APPLICATION NUMBER: US/09/427,261A
;  CURRENT FILING DATE: 1999-10-25
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2193
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:Syndecan DNA
US-09-427-261-3

  Query Match             1.8%;  Score 24;  DB 4;  Length 2193;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       2159 TGTCAAAAAAAAAAAAAAAAAAAA 2182


RESULT 22
US-09-539-333D-222
; Sequence 222, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Bihain, Bernard
;  APPLICANT:  Essioux, Laurent
;  TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
;  FILE REFERENCE: GENSET.047AUS
;  CURRENT APPLICATION NUMBER: US/09/539,333D
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION NUMBER: US 60/126,903
;  PRIOR FILING DATE: 1999-03-30
;  PRIOR APPLICATION NUMBER: US 60/131,971
;  PRIOR FILING DATE: 1999-04-30
;  PRIOR APPLICATION NUMBER: US 60/132,065
;  PRIOR FILING DATE: 1999-04-30
;  PRIOR APPLICATION NUMBER: US 60/143,928
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: US 60/145,915
;  PRIOR FILING DATE: 1999-07-27
;  PRIOR APPLICATION NUMBER: US 60/146,453
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: US 60/146,452
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: US 60/162,288
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: US 09/416,384
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 231
;  SOFTWARE: Patent.pm
; SEQ ID NO 222
;   LENGTH: 3001
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1501
;   OTHER INFORMATION: 99-27349-267 : polymorphic base G or A
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 1502..1521
;   OTHER INFORMATION: 99-27349-267.mis1,  complement
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 1482..1500
;   OTHER INFORMATION: 99-27349-267.mis2
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1748..1767
;   OTHER INFORMATION: upstream amplification primer, complement
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1337..1355
;   OTHER INFORMATION: downstream amplification primer
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 1489..1513
;   OTHER INFORMATION: 99-27349-267  probe
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 182,848,1501,2206,2397
;   OTHER INFORMATION: n=a, g, c or t
US-09-539-333D-222

  Query Match             1.8%;  Score 24;  DB 4;  Length 3001;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       2367 TGTCAAAAAAAAAAAAAAAAAAAA 2390


RESULT 23
US-09-112-096-28
; Sequence 28, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
;  APPLICANT: Reiner Laus
;  APPLICANT: Michael H. Shapero
;  APPLICANT: Larisa Tsavaler
;  TITLE OF INVENTION: Prostate Tumor Polynucleotide and
;  TITLE OF INVENTION: Antigen Compositions
;  FILE REFERENCE: 7636-0015.30
;  CURRENT APPLICATION NUMBER: US/09/112,096
;  CURRENT FILING DATE: 1998-07-09
;  EARLIER APPLICATION NUMBER: 60/056,110
```

```
;  EARLIER FILING DATE: 1997-08-20
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28
;    LENGTH: 3848
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-112-096-28

  Query Match             1.8%;  Score 24;  DB 3;  Length 3848;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db       3780 TGTCAAAAAAAAAAAAAAAAAAAA 3803


RESULT 24
US-08-381-691-16/c
; Sequence 16, Application US/08381691
; Patent No. 5852224
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  Alpha-lac Albumin Gene Constructs
;     NUMBER OF SEQUENCES:  17
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/381,691
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3952 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-381-691-16

  Query Match             1.8%;  Score 24;  DB 2;  Length 3952;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db       3856 TGTCAAAAAAAAAAAAAAAAAAAA 3833


RESULT 25
US-09-112-096-14
; Sequence 14, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
;  APPLICANT: Reiner Laus
;  APPLICANT: Michael H. Shapero
;  APPLICANT: Larisa Tsavaler
;  TITLE OF INVENTION: Prostate Tumor Polynucleotide and
;  TITLE OF INVENTION: Antigen Compositions
;  FILE REFERENCE: 7636-0015.30
;  CURRENT APPLICATION NUMBER: US/09/112,096
;  CURRENT FILING DATE: 1998-07-09
;  EARLIER APPLICATION NUMBER: 60/056,110
;  EARLIER FILING DATE: 1997-08-20
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;    LENGTH: 5668
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-112-096-14

  Query Match             1.8%;  Score 24;  DB 3;  Length 5668;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db       5600 TGTCAAAAAAAAAAAAAAAAAAAA 5623


RESULT 26
US-09-636-215-777
; Sequence 777, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 777
;    LENGTH: 5668
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-636-215-777

  Query Match             1.8%;  Score 24;  DB 4;  Length 5668;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db       5600 TGTCAAAAAAAAAAAAAAAAAAAA 5623


RESULT 27
US-09-685-166A-777
; Sequence 777, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
```

```
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 777
;   LENGTH: 5668
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-777

  Query Match             1.8%;  Score 24;  DB 4;  Length 5668;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db       5600 TGTCAAAAAAAAAAAAAAAAAAAA 5623


RESULT 28
US-09-810-671-3/c
; Sequence 3, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
;  APPLICANT: YAN, Chunhua et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000758
;  CURRENT APPLICATION NUMBER: US/09/810,671
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 21234
;   TYPE: DNA
;   ORGANISM: Human
US-09-810-671-3

  Query Match             1.8%;  Score 24;  DB 4;  Length 21234;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db      20918 TGTCAAAAAAAAAAAAAAAAAAAA 20895


RESULT 29
US-10-109-854-3/c
; Sequence 3, Application US/10109854
; Patent No. 6630337
; GENERAL INFORMATION:
;  APPLICANT: YAN, Chunhua et al.
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000758DIV
;  CURRENT APPLICATION NUMBER: US/10/109,854
;  CURRENT FILING DATE:  2002-04-01
;  PRIOR APPLICATION NUMBER: 60/227,470
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: 09/810,671
;  PRIOR FILING DATE: 2001-03-19
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 21234
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-109-854-3

  Query Match             1.8%;  Score 24;  DB 4;  Length 21234;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db      20918 TGTCAAAAAAAAAAAAAAAAAAAA 20895


RESULT 30
US-09-245-281-44/c
; Sequence 44, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
;  APPLICANT: Bertin, John
;  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
;  TITLE OF INVENTION: AND USES THEREOF
;  FILE REFERENCE: 07334/118001
;  CURRENT APPLICATION NUMBER: US/09/245,281
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: US 09/207,359
;  EARLIER FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 09/099,041
;  EARLIER FILING DATE: 1998-06-17
;  EARLIER APPLICATION NUMBER: US 09/019,942
;  EARLIER FILING DATE: 1998-02-06
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;   LENGTH: 32042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-281-44

  Query Match             1.8%;  Score 24;  DB 4;  Length 32042;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
             ||||||||||||||||||||||||
Db      11653 TGTCAAAAAAAAAAAAAAAAAAAA 11630


RESULT 31
US-09-340-620A-63/c
; Sequence 63, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
;  APPLICANT: Bertin, John
;  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THER
;  FILE REFERENCE: 07334-124001
;  CURRENT APPLICATION NUMBER: US/09/340,620A
;  CURRENT FILING DATE:  1999-06-28
;  PRIOR APPLICATION NUMBER: US 09/245,281
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: US 09/207,359
;  PRIOR FILING DATE: 1998-12-08
;  PRIOR APPLICATION NUMBER: US 09/099,041
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: US 09/019,942
;  PRIOR FILING DATE: 1998-02-06
;  NUMBER OF SEQ ID NOS: 71
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
;   LENGTH: 32042
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-09-340-620A-63

  Query Match             1.8%;  Score 24;  DB 4;  Length 32042;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db      11653 TGTCAAAAAAAAAAAAAAAAAAAA 11630


RESULT 32
US-09-146-053-3/c
; Sequence 3, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
;  APPLICANT: Ryan, James W.
;  APPLICANT: Sprinkle, Terry Joe Curtis
;  APPLICANT: Venema, Richard C.
;  TITLE OF INVENTION: Human Aminopeptidase P Gene
;  FILE REFERENCE: MCG103
;  CURRENT APPLICATION NUMBER: US/09/146,053A
;  CURRENT FILING DATE: 1998-09-02
;  EARLIER APPLICATION NUMBER: 60/057,854
;  EARLIER FILING DATE: 1997-09-02
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 50000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-146-053-3

  Query Match             1.8%;  Score 24;  DB 4;  Length 50000;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db      12947 TGTCAAAAAAAAAAAAAAAAAAAA 12924


RESULT 33
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001067
;  CURRENT APPLICATION NUMBER: US/09/751,389
;  CURRENT FILING DATE:  2001-01-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 786431
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(786431)
;   OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

  Query Match             1.8%;  Score 24;  DB 4;  Length 786431;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||||
Db     537044 TGTCAAAAAAAAAAAAAAAAAAAA 537021


RESULT 34
US-09-306-290-15/c
; Sequence 15, Application US/09306290
; Patent No. 6221635
; GENERAL INFORMATION:
;  APPLICANT: Rovera, Giovanni
;  APPLICANT: Mukhopadhyay, Sunil
;  TITLE OF INVENTION: METHODS FOR SOLID-PHASE AMPLIFICATION OF DNA TEMPLATE
;  TITLE OF INVENTION: (SPADT) USING MULTIARRAYS
;  FILE REFERENCE: 09924-10
;  CURRENT APPLICATION NUMBER: US/09/306,290
;  CURRENT FILING DATE: 1999-05-06
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer RGP
;   OTHER INFORMATION: RR915
US-09-306-290-15

  Query Match             1.7%;  Score 23;  DB 3;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||
Db         23 GTCAAAAAAAAAAAAAAAAAAAA 1


RESULT 35
US-09-621-976-10956
; Sequence 10956, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 10956
;   LENGTH: 67
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-10956

  Query Match             1.7%;  Score 23;  DB 4;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||
Db         32 GTCAAAAAAAAAAAAAAAAAAAA 54


RESULT 36
US-09-621-976-9086
; Sequence 9086, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
```

```
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 9086
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-9086

  Query Match             1.7%;  Score 23;  DB 4;  Length 97;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db         35 GTCAAAAAAAAAAAAAAAAAAA 57


RESULT 37
US-09-091-725-28
; Sequence 28, Application US/09091725
; Patent No. 6329141
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Improved methods for transforming Phaffia
;    TITLE OF INVENTION:  and recombinant DNA for use therein
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster llp
;      STREET:  2000 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  United States of America
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/091,725
;      FILING DATE:  23-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 95203620.0
;      FILING DATE:  22-DEC-1995
;      APPLICATION NUMBER:  EP 96200943.7
;      FILING DATE:  11-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  E. Victor Donahue
;      REGISTRATION NUMBER:  35,492
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  307 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Phaffia rhodozyma
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..227
;      OTHER INFORMATION:  /product= "PRcDNA18"
US-09-091-725-28

  Query Match             1.7%;  Score 23;  DB 4;  Length 307;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db        261 GTCAAAAAAAAAAAAAAAAAAA 283


RESULT 38
US-09-621-976-15342
; Sequence 15342, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 15342
;   LENGTH: 350
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-15342

  Query Match             1.7%;  Score 23;  DB 4;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db        287 GTCAAAAAAAAAAAAAAAAAAA 309


RESULT 39
US-07-987-272A-13
; Sequence 13, Application US/07987272A
; Patent No. 5731166
;  GENERAL INFORMATION:
;    APPLICANT:  Geczy, C., Simpson, R. J. and Lackmann, M
;    TITLE OF INVENTION:  No. 5731166el Chemotactic Factor
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman Darby & Cushman
;      STREET:  1100 New York Avenue, N. W., Ninth Floor, East Tower
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/987,272A
;      FILING DATE:  05-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PK 2127
;      FILING DATE:  05-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PK 4463
;      FILING DATE:  05-SEP-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brinkman, David W
;      REGISTRATION NUMBER:  20,817
;      REFERENCE/DOCKET NUMBER:  DWB/1925/200259
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861 3000
;      TELEFAX:  202-822 0944
;      TELEX:  6714627 CUSH
```

```
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  433 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  52..318
US-07-987-272A-13

  Query Match             1.7%;  Score 23;  DB 1;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db        386 GTCAAAAAAAAAAAAAAAAAAA 408


RESULT 40
US-08-661-168-6
; Sequence 6, Application US/08661168
; Patent No. 5773649
;  GENERAL INFORMATION:
;    APPLICANT:  SINNETT, Daniel
;    APPLICANT:  LABUDA, Damian
;    APPLICANT:  KRAJINOVIC, Maja
;    APPLICANT:  RICHER, Chantal
;    TITLE OF INVENTION:  DNA MARKERS TO DETECT CANCER CELLS
;    TITLE OF INVENTION:  EXPRESSING A MUTATOR PHENOTYPE AND METHOD OF DIAGNOSIS OF
;    TITLE OF INVENTION:  CANCER CELLS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SWABEY OGILVY RENAULT
;      STREET:  1981 McGill College Avenue, Suite 1600
;      CITY:  Montreal
;      STATE:  Quebec
;      COUNTRY:  Canada
;      ZIP:  H3A 2Y3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,168
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COTE, France
;      REGISTRATION NUMBER:  37,037
;      REFERENCE/DOCKET NUMBER:  12667-7US FC/ld
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (514) 845-7126
;      TELEFAX:  (514) 288-8389
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  567 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-661-168-6

  Query Match             1.7%;  Score 23;  DB 1;  Length 567;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db         12 GTCAAAAAAAAAAAAAAAAAAA 34


RESULT 41
US-09-385-982-4
; Sequence 4, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(627)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-4

  Query Match             1.7%;  Score 23;  DB 3;  Length 627;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db        569 GTCAAAAAAAAAAAAAAAAAAA 591


RESULT 42
US-08-874-460-1
; Sequence 1, Application US/08874460
; Patent No. 5981231
;  GENERAL INFORMATION:
;    APPLICANT:  WEI, YING-FEI
;    APPLICANT:  KREIDER, BRENT
;    APPLICANT:  ROSEN, CRAIG
;    TITLE OF INVENTION:  CHEMOKINE BETA 15
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  US
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/874,460
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/019,837
;      FILING DATE:  17-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K.
```

```
;       REGISTRATION NUMBER:  36,688
;       REFERENCE/DOCKET NUMBER:  1488.0420001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-371-2600
;       TELEFAX:  202-371-2540
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  989 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  88..534
;     FEATURE:
;       NAME/KEY:  sig_peptide
;       LOCATION:  88..147
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  148..534
US-08-874-460-1

  Query Match             1.7%;  Score 23;  DB 2;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||
Db        958 GTCAAAAAAAAAAAAAAAAAAAA 980


RESULT 43
US-09-272-162-1
; Sequence 1, Application US/09272162
; Patent No. 6503735
;   GENERAL INFORMATION:
;       APPLICANT: WEI, YING-FEI
;                  KREIDER, BRENT
;                  ROSEN, CRAIG
;       TITLE OF INVENTION: CHEMOKINE BETA 15
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;            STREET: 1100 NEW YORK AVENUE, SUITE 600
;            CITY: WASHINGTON
;            STATE: D.C.
;            COUNTRY: US
;            ZIP: 20005-3934
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/272,162
;            FILING DATE: 19-Mar-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/874,460
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Steffe, Eric K.
;            REGISTRATION NUMBER: 36,688
;            REFERENCE/DOCKET NUMBER: 1488.0420001
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 202-371-2600
;            TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 989 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  88..534
;       FEATURE:
;            NAME/KEY:  sig_peptide
;            LOCATION:  88..147
;       FEATURE:
;            NAME/KEY:  mat_peptide
;            LOCATION:  148..534
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-272-162-1

  Query Match             1.7%;  Score 23;  DB 4;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||
Db        958 GTCAAAAAAAAAAAAAAAAAAAA 980


RESULT 44
US-09-647-224A-5
; Sequence 5, Application US/09647224A
; Patent No. 6482631
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT:  Falco, Saverio Carl
;  APPLICANT:  Gutteridge, Steven
;  APPLICANT:  Hitz, William D.
;  APPLICANT:  Maxwell, Carl A.
;  APPLICANT:  Rafalski, J. Antoni
;  APPLICANT:  Tao, Yong
;  APPLICANT:  Vollmer, Steven J.
;  TITLE OF INVENTION: TRYPTOPHAN BIOSYNTHETIC ENZYMES
;  FILE REFERENCE: BB-1150-A
;  CURRENT APPLICATION NUMBER: US/09/647,224A
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 60/079,386
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: PCT/US99/06046
;  PRIOR FILING DATE: 1999-03-19
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
;   LENGTH: 1005
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-09-647-224A-5

  Query Match             1.7%;  Score 23;  DB 4;  Length 1005;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1303 ATGTCAAAAAAAAAAAAAAAAAA 1325
              |||||||||||||||||||||||
Db        983 ATGTCAAAAAAAAAAAAAAAAAA 1005


RESULT 45
US-09-396-149-17
; Sequence 17, Application US/09396149
; Patent No. 6538176
; GENERAL INFORMATION:
;  APPLICANT: Mahajan, Pramod B.
;  TITLE OF INVENTION: Maize Replication Protein A and Use
;  FILE REFERENCE: 5718-59
;  CURRENT APPLICATION NUMBER: US/09/396,149
;  CURRENT FILING DATE: 1999-09-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 17
;   LENGTH: 1087
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Maize RPA Middle Subunit Homologue-5
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)...(1044)
US-09-396-149-17

  Query Match             1.7%;  Score 23;  DB 4;  Length 1087;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1304 TGTCAAAAAAAAAAAAAAAAAAA 1326
              |||||||||||||||||||||||
Db       1065 TGTCAAAAAAAAAAAAAAAAAAA 1087


RESULT 46
US-08-036-555B-135
; Sequence 135, Application US/08036555B
; Patent No. 5530109
;  GENERAL INFORMATION:
;    APPLICANT:  Goodearl, Andrew; Stroobant, Paul;
;    APPLICANT:  Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
;    APPLICANT:  Chen, Maio Su; Hiles, Ian
;    TITLE OF INVENTION:  Glial Mitogenic Factors, Their
;    TITLE OF INVENTION:  Preparation and Use
;    NUMBER OF SEQUENCES:  184
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/036,555B
;      FILING DATE:  24-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,173
;      FILING DATE:  23-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/940,389
;      FILING DATE:  03-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/907,138
;      FILING DATE:  30-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/863,703
;      FILING DATE:  03-APRIL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.K. 91 07566.3
;      FILING DATE:  10-APRIL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsai, Christine H.
;      REGISTRATION NUMBER:  34,266
;      REFERENCE/DOCKET NUMBER:  LUD 5250.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-036-555B-135

  Query Match             1.7%;  Score 23;  DB 1;  Length 1108;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              |||||||||||||||||||||||
Db       1059 GTCAAAAAAAAAAAAAAAAAAA 1081


RESULT 47
US-08-469-569-135
; Sequence 135, Application US/08469569
; Patent No. 5606032
;  GENERAL INFORMATION:
;    APPLICANT:  Goodearl, Andrew; Stroobant, Paul;
;    APPLICANT:  Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
;    APPLICANT:  Chen, Maio Su; Hiles, Ian
;    TITLE OF INVENTION:  Glial Mitogenic Factors, Their
;    TITLE OF INVENTION:  Preparation and Use
;    NUMBER OF SEQUENCES:  184
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,569
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/036,555
;      FILING DATE:  24-MAR-1993
;      APPLICATION NUMBER:  07/965,173
;      FILING DATE:  23-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/940,389
;      FILING DATE:  03-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/907,138
;      FILING DATE:  30-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/863,703
;      FILING DATE:  03-APRIL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.K. 91 07566.3
;      FILING DATE:  10-APRIL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsai, Christine H.
;      REGISTRATION NUMBER:  34,266
;      REFERENCE/DOCKET NUMBER:  LUD 5250.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-469-569-135
```

```
  Query Match             1.7%;  Score 23;  DB 1;  Length 1108;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db       1059 GTCAAAAAAAAAAAAAAAAAAA 1081


RESULT 48
US-08-249-322A-135
; Sequence 135, Application US/08249322A
; Patent No. 5716930
;  GENERAL INFORMATION:
;    APPLICANT:  Goodearl, Andrew; Stroobant, Paul;
;    APPLICANT:  Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
;    APPLICANT:  Chen, Maio Su; Hiles, Ian
;    TITLE OF INVENTION:  Glial Mitogenic Factors, Their
;    TITLE OF INVENTION:  Preparation and Use
;    NUMBER OF SEQUENCES:  184
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,322A
;      FILING DATE:  26-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/036,555
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,173
;      FILING DATE:  23-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/940,389
;      FILING DATE:  03-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/907,138
;      FILING DATE:  30-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/863,703
;      FILING DATE:  03-APRIL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.K. 91 07566.3
;      FILING DATE:  10-APRIL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsai, Christine H.
;      REGISTRATION NUMBER:  34,266
;      REFERENCE/DOCKET NUMBER:  LUD 250.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-249-322A-135

  Query Match             1.7%;  Score 23;  DB 1;  Length 1108;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db       1059 GTCAAAAAAAAAAAAAAAAAAA 1081


RESULT 49
US-08-469-526A-135
; Sequence 135, Application US/08469526A
; Patent No. 5792849
;  GENERAL INFORMATION:
;    APPLICANT:  Goodearl, Andrew
;    APPLICANT:  Stroobant, Paul
;    APPLICANT:  Minghetti, Luisa
;    APPLICANT:  Waterfield, Michael
;    APPLICANT:  Marchionni, Mark
;    APPLICANT:  Chen, Maio Su
;    APPLICANT:  Hiles, Ian
;    TITLE OF INVENTION:  GLIAL MITOGENIC FACTORS, THEIR
;    TITLE OF INVENTION:  PREPARATION AND USE
;    NUMBER OF SEQUENCES:  187
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Clark & Elbing LLP
;      STREET:  176 Federal Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,526A
;      FILING DATE:  06 June 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/036,555
;      FILING DATE:  24-MAR-1993
;      APPLICATION NUMBER:  07/965,173
;      FILING DATE:  23-OCT-1992
;      APPLICATION NUMBER:  07/940,389
;      FILING DATE:  03-SEP-1992
;      APPLICATION NUMBER:  07/907,138
;      FILING DATE:  03-JUN-1992
;      APPLICATION NUMBER:  07/863,703
;      FILING DATE:  03-APRIL-1992
;      APPLICATION NUMBER:  U.K. 91 07566.3
;      FILING DATE:  10-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bieker-Brady, Kristina
;      REGISTRATION NUMBER:  39,109
;      REFERENCE/DOCKET NUMBER:  04585/00200A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-428-0200
;      TELEFAX:  617-428-7045
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  8...778
;      OTHER INFORMATION:
US-08-469-526A-135

  Query Match             1.7%;  Score 23;  DB 1;  Length 1108;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db       1059 GTCAAAAAAAAAAAAAAAAAAA 1081


RESULT 50
US-08-734-591A-135
; Sequence 135, Application US/08734591A
; Patent No. 5854220
;  GENERAL INFORMATION:
;    APPLICANT:  Goodearl, Andrew
;    APPLICANT:  Stroobant, Paul
;    APPLICANT:  Minghetti, Luisa
;    APPLICANT:  Waterfield, Michael
;    APPLICANT:  Hiles, Ian
;    APPLICANT:  Marchionni, Mark
;    APPLICANT:  Chen, Mario
;    TITLE OF INVENTION:  GLIAL MITOGENIC FACTORS, THEIR
;    TITLE OF INVENTION:  PREPARATION AND USE
;    NUMBER OF SEQUENCES:  187
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Clark & Elbing LLP
;      STREET:  176 Federal Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM Compatible Pentium
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  WordPerfect (Version 7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/734,591A
;      FILING DATE:  22-OCT-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/470,335
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/036,555
;      FILING DATE:  03-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,173
;      FILING DATE:  23-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/940,389
;      FILING DATE:  03-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/907,138
;      FILING DATE:  30-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/863,703
;      FILING DATE:  03-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  UK 91 07566.3
;      FILING DATE:  10-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bieker-Brady, Kristina
;      REGISTRATION NUMBER:  39,109
;      REFERENCE/DOCKET NUMBER:  04585/00200P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 428-0200
;      TELEFAX:  (617) 428-7045
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  8...778
;      OTHER INFORMATION:
US-08-734-591A-135

  Query Match             1.7%;  Score 23;  DB 2;  Length 1108;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1305 GTCAAAAAAAAAAAAAAAAAAA 1327
              ||||||||||||||||||||||
Db       1059 GTCAAAAAAAAAAAAAAAAAAA 1081


Search completed: September 17, 2004, 05:42:46
Job time : 129 secs
```